TAMPERE UNIVERSITY OF TECHNOLOGY
Department of Communications Engineering

SHWETA SHRESTHA

RSS-BASED POSITION ESTIMATION IN CELLULAR AND WLAN NETWORKS

Master of Science Thesis

Topic approved by:
Faculty Council of
Computing and Electrical Engineering on
11th Jan, 2012

Examiners:
Adjunct Professor Dr. Tech Elena- Simona Lohan
Professor Dr. Tech. Mikko Valkama

Abstract

TAMPERE UNIVERSITY OF TECHNOLOGY
Master's Degree Program in Information Technology
SHRESTHA, SHWETA: RSS-based position estimation in cellular and WLAN networks
Master of Science Thesis, Pages, 68
March 2012
Major subject: Communication Engineering
Examiner(s): Adjunct Professor, Dr. Tech. Elena- Simona Lohan
Professor, Dr. Tech. Mikko Valkama
Keywords: Received signal strength (RSS), Wireless local area networks (WLANs), fingerprinting, path-loss model, Access points (APs), Root mean square error (RMSE)

Indoor positioning and tracking is gaining more and more interest in the research field, motivated on one hand by the desire to achieve seamless ubiquitous positioning and, on the other hand, by the potential of huge Location Based Services revenues in the future. The received signal strength (RSS)-based positioning has become a key to enter into the field of indoor and urban positioning where Global Positioning System (GPS) often fails to serve. The main objective of this thesis has been to bring contributions in the RSS-based indoor or urban positioning by implementing a generic multi-purpose MATLAB simulator.

The motivation of such a work has been partly from literatures and partly from our experimental findings, which shows that RSS-based positioning often requires time consuming measurement campaigns that are site dependent. The built indoor multi-floor MATLAB simulator incorporates adjustable parameters, based on the analysis and statistics from performed measurements from both cellular and WLAN signals. Then accordingly it generates new scenarios and propagation effects and then allows us to output the user positioning accuracy in terms of Root Mean Square Errors. The additional aim is to analyze in more comprehensive way different RSS-based positioning methods. Typically, the current published research work based on RSS-based positioning is limited to fingerprinting methods. Both the fingerprinting and the path-loss method using the Euclidian distance approaches have been implemented in positioning the user in this thesis work. Apart from user estimation the estimation of the AP location and the path-loss parameters are also carried out in this thesis.

The thesis results contain two parts: one based on measured data and another based on simulation. The positioning accuracy as estimated via Monte-Carlo simulations with the built simulator is compared with the results based on real measured data. A high correlation between the two has been observed, as expected with the Monte-Carlo accuracy results coming from the simulations being slightly worse than those observed in location with real data, because a larger set of scenarios could be captured with our simulator.

Preface

This Master of Science Thesis has been carried out in the Department of Communication Engineering (DCE) at Tampere University of Technology (TUT), Tampere, Finland as a part of Nokia's Rx Positioning project during June 2011 – Jan 2012. The thesis work has been jointly funded by two ongoing research projects at the Department: the Academy of Finland-funded project "Digital Processing Algorithms for Indoor Positioning Systems (ACAPO)" and Rx Positioning project, Nokia, Tampere, Finland which are greatly acknowledged.

I am pleased to express my gratitude to my thesis supervisors, Adjunct Professor Dr. Elena-Simona Lohan and Professor, Dr. Mikko Valkama, for their valuable guidance and support throughout the thesis period. I would also like to express my appreciation to my seniors Elina Laitinen and Jukka Talvitie, who have worked in the same project. Their way of sharing knowledge and willing to help attitude has supported me a lot to foster my work in the correct direction. I am also grateful to Jari Syrjärinne, Lauri Wirola, Mikko Blomqvist and Tommi Laine from Nokia Company for their valuable comments and suggestions regarding our results.

Finally, it is a pleasure to thank my family and friends for their support and motivation to pursue this Master's degree.

Tampere, 20th Feb, 2012

Shweta Shrestha

Table of Contents

List of Abbreviations

AOA	Angle of Arrival
AP	Access Point
BS	Base Station
dB	Decibel
dBm	Decibel Miliwatt
DCE	Department of Communication Engineering
GNSS	Global Navigation Satellite Systems
GPS	Global Positioning System
GSM	Global System for Mobile communications
GUI	Graphical User Interface
LBS	Location Based Services
LOS	Line-of-Sight
MAC	Media Access Control
MBWA	Mobile Broadband Wireless Access
MIMO	multiple-input and multiple-output
MS	Mobile Station
NLOS	Non Line-of-Sight
NN	Nearest Neighbor
PDF	Probability Density Function
RFID	Radio Frequency Identification
RMSE	Root Mean Square Error

RSS	Received Signal Strength
RSSI	Received Signal Strength Indicator
SSID	Service Set Identifier
TDOA	Time Difference of Arrival
TOA	Time of Arrival
TUT	Tampere University of Technology
WCDMA	Wideband Code Division Multiple Access
WiBro	Wireless Broadband
Wi-Fi	Wireless Fidelity
WiMax	Microwave Access
WLAN	Wireless Local Area Network
WMAN	Wireless Metropolitan Area Network
WPAN	Wireless Personal Area Network
WRAN	Wireless Regional Area Network
WWAN	Wireless Wide Area Network

List of Symbols

P_T	Transmit power of Access point/ Base Station
n	Path-loss exponent
Ψ	Shadowing deviation
σ	Shadowing standard deviation
μ	Mean shadowing
c	velocity of light in free space
λ	wavelength
f	frequency

Table of Tables

Table of Figures

1. Introduction

1.1. Background and Introduction

In this moving era of wireless and mobile communication, seamless connectivity, accessibility and portability is essential anywhere, everywhere. No matter wherever a mobile user happens to be, he/she wants to have access to right information at the right time [1]. This brings in the need of location tracking or user positioning. The use of location tracking and navigation started with civilian navigation but now it is acquiring global use in the general public as well. In the commercial field satellite positioning started with the traditional use of Global Positioning System (GPS) in transportation. Nowadays there has been a quite extensive penetration of wireless positioning in robotics, security alert systems and smart home appliances through the use of Radio Frequency Identifications (RFIDs) sensors and other sensor integrated systems [2]. With the continuous decrease in size and price of GPS receiver [2] and also with the integration of GPS receiver in mobile phones, the wireless positioning has gained a wide market. For example, according to [3], the total enabled Global Navigation Satellite Systems (GNSS) market size in 2020 is estimated at EUR 244 billion; the core global GNSS market is estimated at EUR 165 billion. Additionally, annual core GNSS revenues in Location Based Services (LBS) are expected to grow between 2010 and 2020 from EUR 12 billion to EUR 96 billion [3]. But the shortcomings of satellite navigation cannot be overlooked. In indoor locations and urban environments, GPS is not efficient as the signals are too weak and suffer severe multipath and interference degradation [2], [4]. So, the GPS signals cannot be relied on for indoor and urban positioning.

The shortcomings of GPS signals in urban and indoor scenarios bring the new innovative areas to research for in the field of wireless communication. As a result, it requires alternative methods besides GPS signals for indoor or urban positioning. Recently, there has been a widespread deployment of Wireless Local Area Networks (WLANs) in many offices, campuses and many individual's homes and organizations [1]. Also the ubiquitous accessible cellular network is gaining its popularity with the system being upgraded and extended every time. So, why not to go with those existing wireless systems such as cellular systems and WLANs? Instead of implementing a specialized infrastructure only for positioning, in order to overcome the shortcomings in indoors and urban areas, the existing wireless systems can be employed [1]. The Received Signal Strength (RSS) from easily accessible infrastructure such as WLANs and cellular networks can be used for positioning in those areas where the GPS satellite positioning fails to serve. Thus, due to the ease of

accessibility, no more requirements of extra cost and additional system, more and more demand and supply of those wireless communication network infrastructure have attracted the researchers to focus their interests towards the issues of RSS-based positioning using the existing WLAN and cellular systems [1], [2], [5]. WLAN RSS-based positioning is growing rapidly in importance [24], [25]. RSS-based localization systems use the RSS from the available Access Points (APs)/ Base Stations (BSs) to determine the location of the Mobile Station (MS) [27], [28].

Table 1.1 *Wireless network standards.*

Wireless area networks	**Typical operational range**	**Operating Frequency bands/carriers**	**Examples**
Wireless Personal Area Networks (WPANs)	0-10 m	2.4GHz	Bluetooth [7], ZigBee [8], IEEE 802.15 standards [8], [7]
Wireless Local Area Networks (WLANs)	0-150 m	2.4GHz, 5GHz	IEEE 802.11 standards [7] , Wireless Fidelity (Wi-Fi)
Wireless Metropolitan Area Networks (WMANs)	0-5 km	2-11GHz [8]	IEEE 802.16 standards [7], Worldwide Interoperability for Microwave Access (WiMax) [8], Wireless Broadband (WiBro)
Wireless Wide Area Networks (WWANs)	0-15 km	Less than 3.5GHz	IEEE 802.20 standards, Mobile Broadband Wireless Access *(*MBWA) [7], [8]
Wireless Regional Area Networks (WRANs)	0-100 km	54-862MHz [6]	IEEE 802.22 standards [6],[9]

In *Table 1.1*different wireless network standards, their operating frequency and range and the example technologies are shown. The wireless area networks measured in our thesis are most likely to belong to 802.11 family, since WLANs belonging to this family are usually deployed in indoor areas. However, we usually have not checked to which standard the

received signal belongs and we used only the RSS values. As seen in *Table 1.1*, various standards may operate at various carrier frequencies. This fact has not been modeled explicitly in our models, but it appeared inherently as a factor modifying the apparent transmit powers and possibly also the path-loss coefficient associated with a certain AP.

In *Table 1.2* [9] different mobile cellular network generations and the technologies used under each generation have been shown. So far in this thesis we have considered only the 2G Global system for mobile communications (GSM) and 3G Wideband Code Division Multiple Access (WCDMA) cellular networks.

Table 1.2 *Mobile cellular network standards.*

Mobile Cellular Network generation	**Technology**
1G	AMPS, TACS
2G	GSM, CDMA
3G	WCDMA, CDMA-2000
4G	LTE

1.2. Thesis objectives

This thesis addresses the problem of cellular and WLAN-based positioning using RSS measurements. The objectives have been:

1. Understanding the propagation channel characteristics (path-losses and shadowing) based on WLAN and cellular measured RSS values
2. Modeling as accurately as possible the realistic channel (e.g., investigate path-loss parameters, distribution of RSS, power maps, shadowing characteristics, etc)
3. Building a MATLAB simulator for indoor WLAN positioning using some of the parameters extracted from the measured channels
4. Investigating the performance (e.g., based on Root Mean Square Error (RMSE)) in locating a Mobile Station (MS) with both measured and simulated data

The main novel contributions of the Author have been at Objectives 2, 3 and 4. The author has also been main author and co-author in 2 accepted conference papers [4], [10].

1.3. Thesis organization

The rest of the thesis is organized as follows:

Chapter 2 briefly describes about the different path-loss models, shadowing and fading that the real radio propagation channel experiences.

The different methods used in positioning with a bit more focus on the main theme of this thesis, 'RSS-based positioning' is presented in Chapter 3.

Chapter 4 presents some real-time channel measurements and their analysis. This shows the channel shadowing or fading effect in terms of radio maps, shadowing standard deviation and some estimated path-loss exponent n and apparent transmit power P_T for indoor and outdoor WLAN and cellular networks based on some real true measurements.

In Chapter 5, there is an overview of a new MATLAB simulator built by the Author, which tries to emulate the real scenarios for the RSS-based location estimation.

The simulation results from the MATLAB simulator developed are shown in the Chapter 6.

Conclusions and future works are addressed in the final Chapter 7.

2. Wireless channel impairments

Unlike the wired channels, wireless radio channels place lots of challenges in the transmission of radio signals from transmitter to receiver. The transmission path between the transmitter and receiver can vary from simple line-of-sight (LOS) to one that is severely obstructed by buildings, mountains and foliage [11]. The wireless radio channels are extremely random and unpredictable. Modeling the radio channels has been one of the most difficult parts of mobile radio system design [11] and is often modeled using the simple free space model or in statistical or empirical way. This thesis will focus on the path-loss and shadowing modeling of cellular and WLAN channels, as needed further in RSS-based positioning.

There are three main losses in the radio path that a received signal experiences on its propagation path: [12]

- Range dependent loss (path-loss)
- Shadowing/ Shadow fading/ Long-term fading
- Fading/ Fast fading /short-term fading

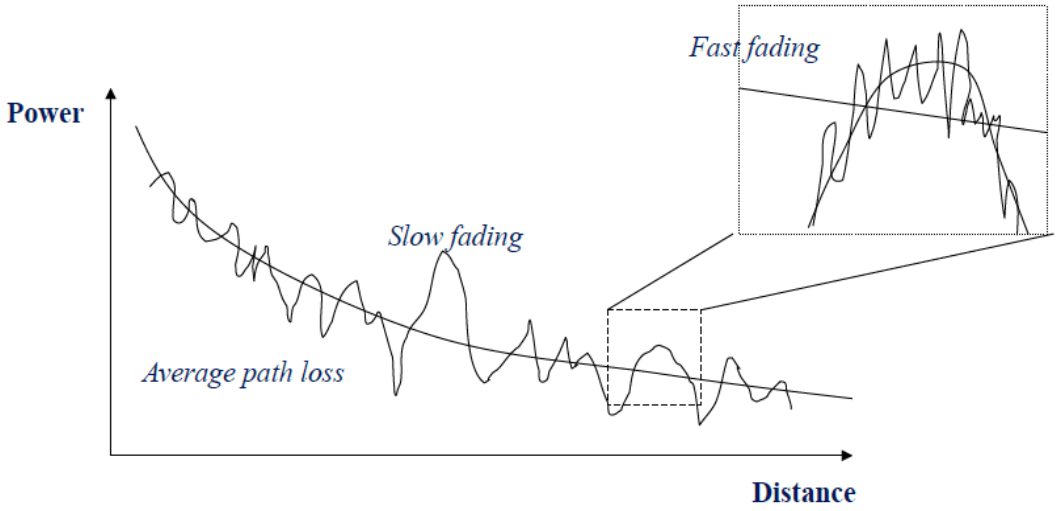

Figure 2.1 *Different attenuations seen in a received signal in a radio channel.*

*Figure 2.1*shows the different kinds of impairments that a signal experiences in its radio path. The next sections give a brief overview of various wireless channel impairments and their typical models.

2.1. Range dependent loss/ Path-loss

The mean received signal or RSS decreases as we move away from the transmitter. According to the free space loss, the RSS is inversely proportional to the square of the distance between the transmitter and the receiver. The path-loss represents the signal attenuation measured in dB between the effective transmitted power and the received power [11]. Different propagation models have been introduced for modeling the path-loss. Some of the widely used path-loss propagation models are summarized in the next sub-sections.

2.1.1. Free space path-loss model

This model predicts the RSS when there is clear LOS path between the transmitter and the receiver. It simply models the received power as a function of transmitter – receiver separation distance raised by power 2 .The free space received power, in linear scale, is given by the Friis equation as in (2.1)[11]:

$$P_R = P_T G_T G_R \left[\frac{\lambda}{4\pi d}\right]^2 \quad (2.1)$$

where, P_R is the received power at the receiver, P_T is the transmitted power, G_T is the transmitter antenna gain, G_R is the receiver antenna gain, λ is the wavelength of the transmitted signal and d is the transmitter – receiver separation distance.

Thus the free space path-loss in decibel (dB) is given as equation (2.2):

$$L_{dB} = 10log_{10}\left[\frac{4\pi d}{\lambda}\right]^2 = 32.4 + 20log_{10}d_{km} + 20log_{10}f_{MHz} \quad (2.2)$$

which comes from c = λ/f = velocity of light and the used units.

2.1.2. Log – distance path-loss model / Simplified path-loss model

This model predicts that the RSS decreases logarithmically with distance. The received power, in linear scale, is given by equation (2.3) [11]:

$$P_R(d) = P_R(d_0)\left[\frac{d_0}{d}\right]^n \quad (2.3)$$

where, $P_R(d_0)$ is the received power at the reference point d_0 which is close to the transmitter (e.g., 1 m), d is the distance from the reference point and n is the path-loss exponent. The path-loss in dB is given by equation (2.4):

$$L_{dB} = 10nlog_{10}\left[\frac{d}{d_0}\right] \tag{2.4}$$

The value of the path-loss coefficient *n* depends on the propagation environment. The value of *n* is 2 for free space and may have a larger value in the environments with obstacles. Below, there is *Table 2.1* which shows some typical values of *n* in different environments, as reported in literature [11]. However, these are only some limiting examples and large variations from these reported values may be and have been observed in field measured data during our measurement analysis.

Table 2.1 *Path-loss exponents for different environments.*

Environment	**Path-Loss Exponent, *n***
Free space	2
Urban area cellular radio	2.7 to 3.5
Shadowed urban cellular radio	3 to 5
In building LOS	1.6 to 1.8
Obstructed in building	4 to 6
Obstructed in factories	2 to 3

All the simulations and the results in this thesis work are based on this simplified path-loss model due to its small number of parameters to be estimated and to its applicability to various scenarios (both indoors and outdoors).

2.1.3. Okumura-Hata model and COST 231 model

This propagation model is an empirical model used to predict the coverage of cellular networks in macro cells in urban and sub urban environments. Okumura-Hata model is valid from 150 till 1500 MHz and it was later extended to COST 231 model which is valid

from 1500 MHz to 2GHz. The path-loss according to Okumura Hata/ COST 231 model is given by equation (2.5) [11]:

$$L_{dB} = A + Blog_{10}(f_{MHz}) - 13.82log_{10}(h_b) + (c - 6.55log_{10}(h_b))log_{10}(d_{km}) - k \quad (2.5)$$

where,

h_b - Base station antenna height [m]

d_{km} - Distance between transmitter and receiver [km]

f_{MHz} - Carrier frequency [MHz]

c - Tunable parameter [44-47]

K - Correction factor [default = 0]

and A and B are frequency dependent parameters given as in *Table 2.2*:

Table 2.2 *A and B parameters in Okumura Hata and COST 231 model.*

	150-1000 MHz	**1500-2000 MHz**
A	69.55	46.3
B	26.16	33.9

Still, such models do not cover well the WLAN spectra, which start from 2.4 GHz to tens of GHz and they also have quite many parameters that make them less suitable for generic RSS-based positioning.

2.1.4. Floor and wall factor model

This is also an empirical model meant for the indoor propagation model. This model takes into account the attenuation due to number of floors and walls in the propagation path. The path-loss according to this model is given as equation (2.6):

$$L_{dB} = L_{ref} + 20log_{10}d + n_f a_f + n_w a_w \quad (2.6)$$

where, n_f and n_w are the number of floors and number of walls, respectively, a_f and a_w are the floor attenuation factor and wall attenuation factor respectively, L_{ref} is the reference

path-loss at d=1 m distance and d is distance between transmitter and receiver in m. Above, it was assumed that we have equal attenuation per floor and per wall inside a building. This assumption is not always true. *Table 2.3* shows some floor and wall attenuations at various operating frequencies due to common building materials as found in literature [11].

Table 2.3 *Attenuation due to common building materials [11].*

	Loss (dB)	**Frequency**
All metal	26	815 MHz
Concrete block wall	8 – 15	1300 MHz
Concrete floor	10	1300 MHz
Dry plywood (3/4 in) – 2 sheets	4	9.6GHz

2.1.5. ITU-R model

This model has similar approach as *Floor and wall factor model* but it explicitly takes the floor attenuation into account. The loss between points on same floor is included implicitly by changing path-loss exponent. According to this model, the path-loss L_{dB} is given as in equation (2.7). [12]

$$L_{dB} = L_{ref} + 20log_{10}f_c + 10nlog_{10}d + L_f(n_f) - 28 \quad (2.7)$$

where, L_{ref} is the reference path-loss at d =1 m distance, f_c is the transmitted frequency in MHz, d is the distance between transmitter and receiver in m and $L_f(n_f)$ is the floor penetration factor.

2.2. Shadowing/ Shadow fading/Long-term fading

In the real radio channels, there are always some obstacles in the propagation path of the radio signals. In the outdoor environment, those obstacles may be buildings, hills, trees etc., while in the indoor environment, the walls, the floors, or the furniture may be some of those obstacles that prevent the direct LOS propagation of radio signals between the transmitter and the receiver. The attenuation in the RSS due to the obstruction of direct LOS path due to the some static obstacle in the propagation path is called shadowing/ shadow fading/ long-term fading.

The above-mentioned path-loss models consider only the variation in mean signal strength due to distance between transmitter and receiver. But in real time modeling, the mean

signal strength might vary vastly than it is predicted from the path-loss models. For a moving MS the RSS is not just a term dependent on the distance from transmitter (as predicted by simple propagation models). But the signal level is highly affected by shadowing from large obstacles like hills, buildings along the signal path. The shadowing is the variation of local mean signal level over a wide area and the local mean is the mean value of fast fading component [12]. Since the envelope of shadowing typically follows a log-normal distribution in linear scale, it is also called log-normal fading [12]. It is generally characterized by using Gaussian or normal distribution in logarithmic scale with zero mean and certain standard deviation. The probability density function (PDF) shadow fading is given as equation (2.8) [13]:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{(x-\mu)^2}{(2\sigma^2}\right)} \tag{2.8}$$

where x is the received shadow fading signal, μ its mean value and σ its standard deviation.

Now if we add shadowing in the above simplified path-loss model's equation, then the combined path-loss and shadowing model in linear scale is given by the equation (2.9) [11]:

$$P_R\,(d) = P_R\,(d_0)\left(\frac{d_0}{d}\right)^n \Psi \tag{2.9}$$

and in dB scale:

$$P_R\,(d) = P_R\,(d_0) - 10nlog_{10}\left(\frac{d}{d_0}\right) + \Psi_{dB} \tag{2.10}$$

where, Ψ_{dB} is typically modeled as a zero mean Gaussian distributed random variable (in dB) with certain standard deviation σ (also in dB).

2.3. Fading/ small-scale fading/ fast fading

Fading simply describes the rapid fluctuation of mean signal strength over short duration of time over a short area [11]. The rapid fluctuation is caused due to the interference between two or more version of same transmitted signals, which arrives at the receiver through different paths at slightly different time. Hence, fading is characterized by the rapid change in amplitude and phases of the received signals caused by multipath propagation. Factors

influencing the fading are: multipath propagation, relative mobility of the receiver and transmitter, transmission bandwidth of the signal [11]. Some of the statistical models used to model the fast fading are Rayleigh, Ricean, Nakagami for the amplitude variations and Clarke's distribution for the spectrum modeling [11].

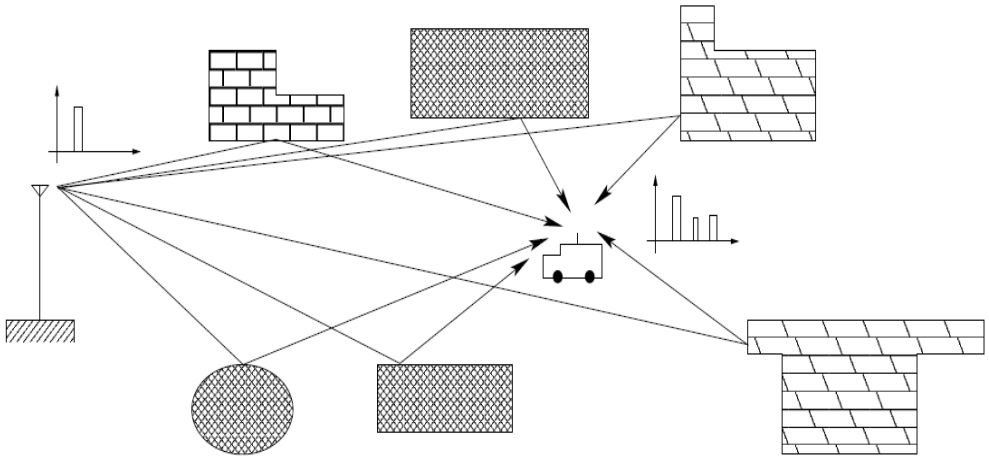

Figure 2.2 *Fading due to multipath propagation.*

Figure 2.2 shows the multiple versions of the transmitted signal that the receiver receives due to multipath.

3. Principles of RSS-based location methods

Basically, there are two steps involved in estimating the location: the first step is the parameter estimation stage, which can be time of arrival (TOA) estimation, time difference of arrival (TDOA) estimation, angle of arrival (AOA) estimation or signal power/ RSS estimation and the second step is the actual location estimation [4]. The location estimation step involves the processing of estimated parameter using different range based or range-free algorithms [14] such as trilateration, proximity or fingerprinting [4].

3.1. Parameter estimation stage

The main idea here is to estimate the distances between the transmitters and receiver. The distance estimation can be implemented in many ways, as explained in the following paragraphs.

1. Time of arrival (TOA):

The distance between the receiver and transmitter is calculated from the propagation time of the transmitted electromagnetic signals, assuming that the speed of propagation (i.e. the speed of light) is known in advance. The distance is calculated using equation (3.1).

$$distance = c * TOA \tag{3.1}$$

where, c is the velocity of light in free space and TOA is the time of flight of the transmitted signal. Though this localization technique provides high accuracy [14] in terms of position estimation, it requires strict time synchronization between the transmitter and receiver [15]. Moreover, TOA algorithms require knowledge about the signal structure (e.g., modulation, multiple access) and are not well suited for WLANs for example, where both CDMA and OFDM techniques are encountered and where modulations range from BPSK to 64-QAM.

2. Time difference of arrival (TDOA):

In this technique, a single receiver receives the signal transmitted from multiple transmitters and, according to the difference of time of arrival of signals, it provides the

distances between the transmitters. Thus, this also requires some synchronization between receiver and transmitter. The TDOA principle is illustrated in *Figure 3.1* below.

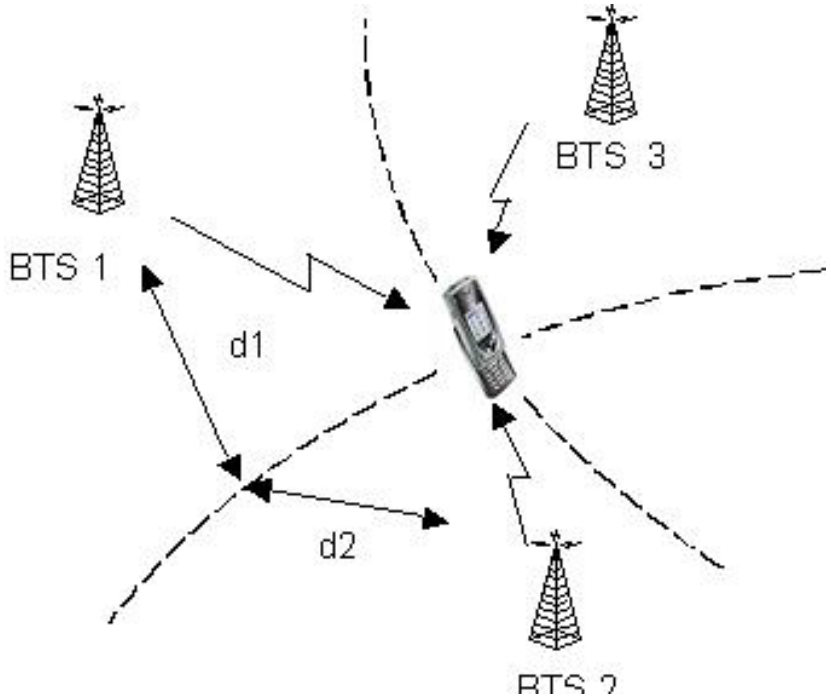

Figure 3.1 *Illustration of TDOA method in cellular case.*

The distance between the transmitters is estimated as in equation (3.2):

$$(d_1 - d_2) = c * TDOA \ = c * (TOA1 - TOA2) \tag{3.2}$$

where, d_1 is the distance between the receiver and first transmitter, d_2 is the distance between the receiver and second transmitter, c is the velocity of light in free space, $TOA1\ and\ TOA2$ are the time of arrival of signal from first and second transmitter respectively. The measured quantity is in fact $TDOA = TOA1 - \ TOA2$.

3. Angle of arrival (AOA):

This technique uses the direction of angle of arrival of radio waves to estimate the MS. At least two AOA measurements or an AOA measurement with TOA or RSS measurement is used to estimate the distance and then the MS location [16]. This requires complex hardware to measure the incidence angle [15].

4. Received signal strength (RSS):

RSS values represent the received power in decibel (dB) or decibel miliwatt (dBm) in any wireless devices (mobile phones, PDAs, wireless network adapters or any other devices equipped with WLAN interfaces). RSS is the strength of incoming signal in a receiver. Strictly speaking each receiver measures a Received Signal Strength Indicator (RSSI), manufacturer-dependent, that is then mapped to an RSS value (in dBm).

Generally, higher RSS value is yielded from a strong signal. The RSS is proportional to the transmitted power and inversely proportional to the square (in free space) or some power (in other environments) of the distance to the source. This propagation law is the basis for estimating distance and location in RSS-based trilateration method.

RSS is a single such parameter which can be used in two ways to estimate MS location. It can be used as a parameter in range based algorithm such as triangulation/trilateration in order to find the distance between the transmitter and receiver. Also, it can be used as a parameter for range free database comparison algorithms such as fingerprinting or some path-loss models derived from the measured RSS database. Much attention has been paid to WLAN- RSS-based indoor positioning algorithm these days [29]. The RSS-based estimation is what is used in this thesis.

3.2. Location estimation stage

3.2.1. RSS-based trilateration

According to the propagation law, the received signal strength at a receiver decreases as distance from the transmitter increases. If the relationship between the RSS and distance is known analytically or empirically, the distance between the two terminals can be determined [15]. When we have a receiver (MS) that can hear three or more transmitters, trilateration can be applied to determine the MS location [15].

Trilateration refers to positioning an object based on the measured distances between the object and multiple reference points at known positions [17]. For example, the traditional GPS uses TOA system plus trilateration principle. Trilateration principle can be in the TOA/TDOA localization method as well as in RSS-based location estimation according to the estimated parameter in the training phase. Since the thesis is based on RSS-based estimation, our concern is over the RSS-based trilateration only.

If the AP/BS location (x_{ap}, y_{ap}, z_{ap}) is known or estimated in advance, and the unknown MS location is denoted via (x, y, z) then,

$$d_{ap} = \sqrt{(x - x_{ap})^2 + (y - y_{ap})^2 + (z - z_{ap})^2} \tag{3.3}$$

gives the distance between the MS and the AP/BS. If we have three such equations from three different APs/BSs the unknown MS location (x,y,z) can be estimated solving the three equations, provided that there is no any errors. With only one hearable transmitter (AP/BS), the uncertainty of 2D location will be on a circle with radius d_{ap} as shown in *Figure 3.2*, with two hearable transmitters, the uncertainty of 2D location is reduced to two points i.e. the intersection of two circles as in *Figure 3.3* and with 3 or more transmitters, we can estimate exactly the (*x*,*y*,*z*) location of MS, in the absence of any error, which is shown in the *Figure 3.4*. The figures can be extended to 3D in a straightforward way.

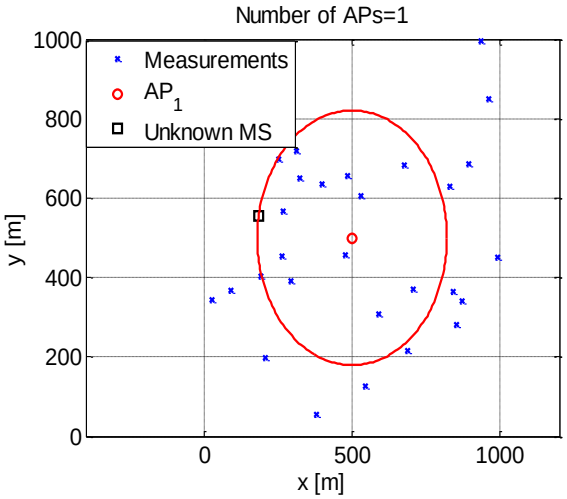

Figure 3.2 *Illustration of uncertainty when one AP is hearable.*

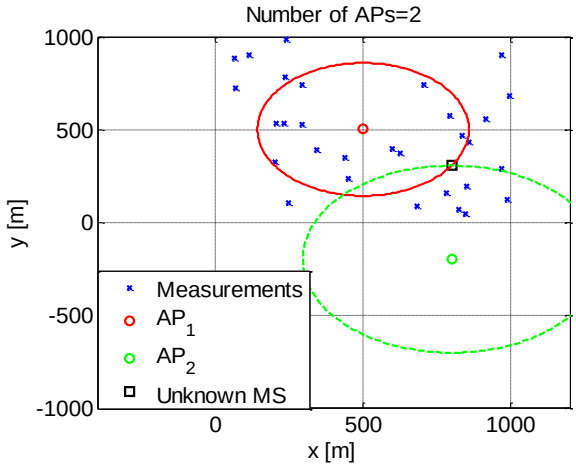

Figure 3.3 *Illustration of uncertainty when two APs are hearable.*

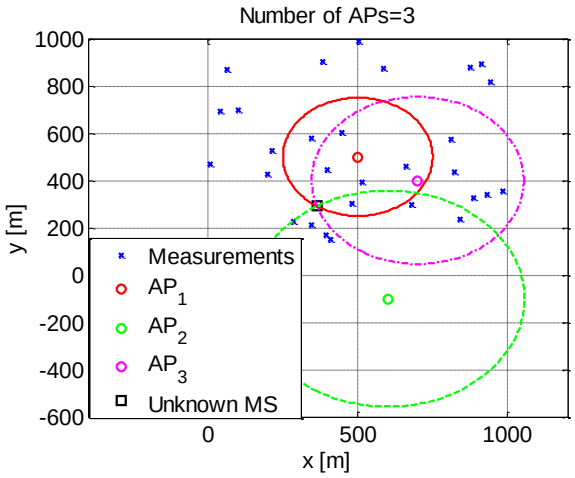

Figure 3.4 *Illustration of uncertainty when three APs are hearable.*

The main difficulty with the RSS-based trilateration method is obtaining the true distance between the MS and the AP/BS. Indoor radio signal propagation is very complicated, because of signal attenuation due to distance, losses through walls and floors and interference due to other electronic devices [18]. The location of AP also may not be known beforehand in all the WLAN cases. Hence, it is difficult to obtain the distance measurement accurately.

3.2.2. RSS-based fingerprinting method

In this technique, there is no any distance measurement between the receiver and transmitter. Instead, the RSS of the receiver is compared to a dataset of RSS taken previously off-line at known locations throughout the coverage area. The reference point of best match in terms of RSS is estimated as the MS location. This type of technique is known as Fingerprinting. Fingerprinting is becoming popular in indoor location estimation. There are two stages in this location estimation method:

- Training or calibration phase
- Positioning phase

Both phases are illustrated in *Figure 3.5* below.

Training phase (data collection/calibration):

The initial training phase is the data collection phase. During this phase, the RSS readings are taken from different reference points in the experimental site. The measurement at those reference points contain the (x,y) (or (x,y,z) in 3D) position of the measured reference point and the vector of RSS values from multiple APs that are heard at that particular reference point. The vector of RSS values at a point is called the location fingerprint of that point [19]. The result of this phase is a database that maps the RSS pattern with the location in that area and this RSS database is called a radio map [19], [20]. It is simply the learning phase of distribution of RSS from each AP over an area in order to estimate the user location in the positioning phase.

Positioning-phase:

The second phase is the positioning phase, in which a positioning algorithm is used to determine the user location based on the RSS at the user point. The RSS at the user point is compared with the RSS database and determined to be that point in the database which find the closest match with the user in terms of RSS [19]. The closest approximation can be determined as the location fingerprint which has minimum difference in RSS with the user point. The minimum difference in RSS can be found using different metrics such as the Euclidean distance, which is the most common distance based technique [19]. In this

algorithm, the Euclidean distance between the user RSS vector and each fingerprint in the database is computed and the position of the fingerprint that provides the smallest Euclidean distance is returned as the estimated position [21]. The closest match can be a single location fingerprint sample or average over 3-5 nearest neighbors (NN). When the NN averaging is used, the MS location is determined by averaging over N_{neigh} closest location fingerprint samples.

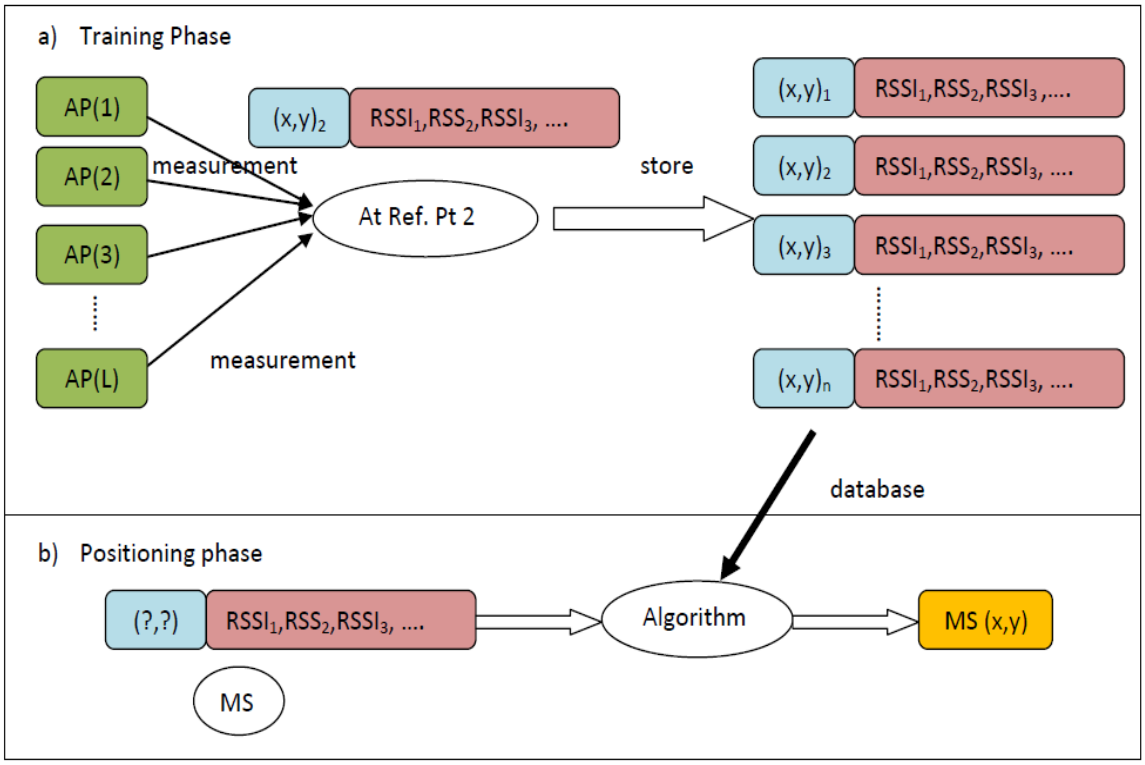

Figure 3.5 *Training and positioning phases in fingerprinting method [18].*

3.2.3. RSS-based path-loss method

Another possible positioning method used with RSS database is the path-loss model. However, this database comparison technique requires huge database collection before the estimation phase. Instead of basing the comparison totally on the actual measurements carried out at each reference positions, some path-loss models can be deployed whose parameters are estimated by fewer samples measurements over the coverage area. [16] This method also has training phase and the estimation phase. *Figure 3.6* shows the illustration of RSS-based path-loss method used in this thesis.

Training phase:

The training phase is the data collection phase similar as in the fingerprinting method.

However it can have fewer samples than in fingerprinting method, since we do not directly compare RSS levels of MS to the RSS levels of the fingerprint database. Apart from this, in the training phase, a path-loss model is built based on the data collected previously in the off-line stage which connects the distance values to the RSS values via a certain path-loss model. The simplest path-loss model used in this thesis is as given below:

$$P_{i,ap} = P_{T,ap} - 10nlog_{10}(d_{i,ap}) + \Psi_i \quad (3.4)$$

where $P_{i,ap}$ is the RSS (in dBm) measured from the ap^{th} AP/BS to the i^{th} measurement point, $P_{T,ap}$ is the transmit power of the ap^{th} transmitter, $d_{i,ap}$ is the distance between the ap^{th} transmitter and i^{th} measurement point, n is a path-loss coefficient characterizing the measured environment, Ψ_i is a noise term including measurement noise and shadowing and fading fluctuations. The unknowns are the (x, y, z) coordinates of the MS, related to the distance to the transmitter via equation (3.3).

In the training phase based on the measured RSS database, the unknown parameters $P_{T,ap}$, n can be determined using some linear regression. And also the AP /BS location (x_{ap}, y_{ap}, z_{ap}) which is unknown in many environments can be also estimated in the training phase, for example by taking an average over the positions of the measurements with the highest RSS.

Positioning phase

Once the path-loss parameters and APs location are estimated in the training phase, the user location can be estimated using different approaches/algorithms such as the Bayesian approach [16] or also recreating own RSS database based on the estimated parameters. So far in this thesis, the MS estimation using the reconstructed RSS grid is only considered.

Once we have the APs location, and the path-loss parameters associated with those APs, we can create a the path-loss grid of uniform grid size and recalculate the RSS values at those grid points using the same path-loss model expressed above in equation (3.4). Now once we have the RSS database we can do the position estimation similarly as above in fingerprinting method.

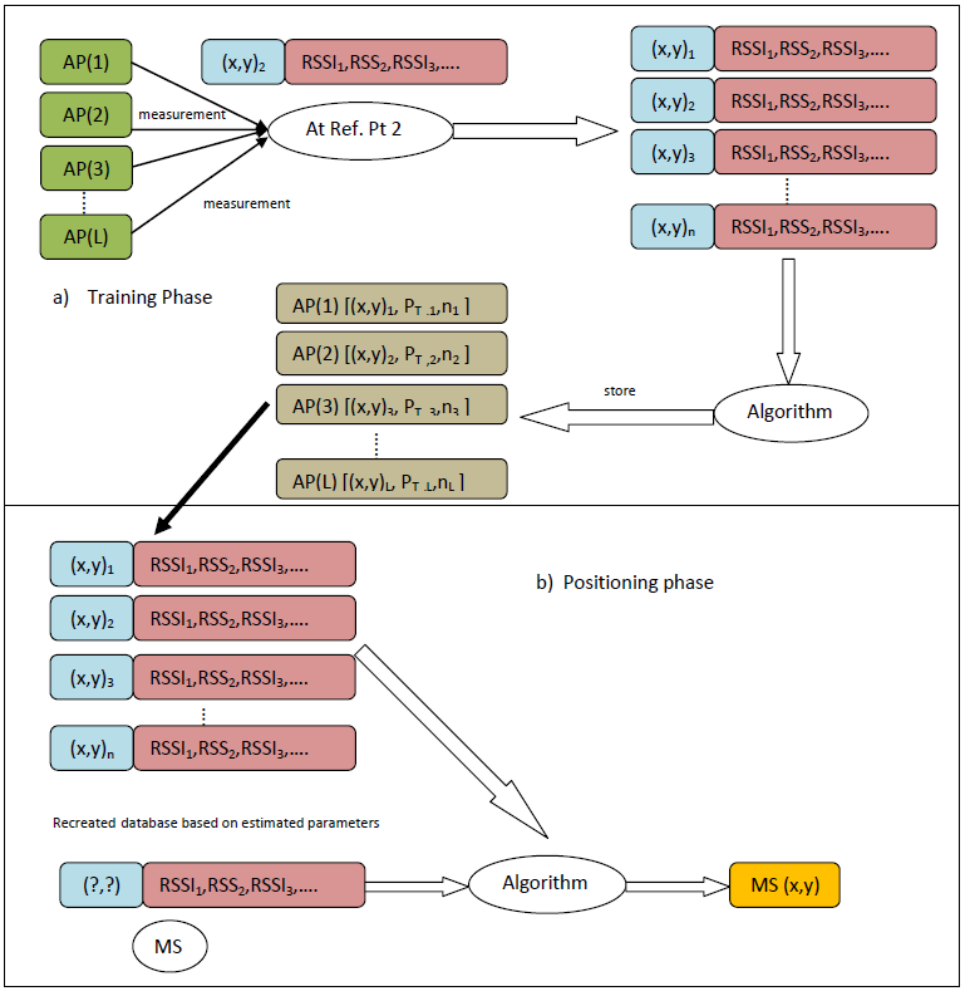

Figure 3.6 *Training and positioning phases in path-loss model.*

3.3. Comparison between RSS-based fingerprinting method and path-loss method

In both fingerprinting and path-loss method, we collect the RSS database in the training phase. In fingerprinting method, there is the direct comparison of RSS levels of MS to the RSS levels of the location fingerprint database. But in the case of path-loss method, from the RSS levels of the location fingerprint database, the parameters like$P_{T,ap}$, n and also the location of APs are determined. Based on the estimated parameters we can recreate our own RSS grid over the coverage area using the above equations (3.3) and (3.4). Then the comparison of the RSS levels of the MS with the RSS levels of the reconstructed path-loss grid is done. Both the location fingerprinting method and path-loss method requires a database that is applicable only for the particular site where it was created, and physical

changes that affect the radio propagation at the site may require recreating a new database. But however, for the path-loss model instead of collecting database at each reference points in the whole coverage area, the path-loss parameters can be estimated over fewer sample measurements over the coverage area. This reduces the burden of storing heavy database as only the path-loss parameters can be stored instead of whole database. Hence, it is quite useful in terms of saving memory in the mobile devices and also the data transfer time and band width in the server side if the database happened to be stored in the server side.

The above discussion is summarized in the following table:

Table 3.1 *Comparison between fingerprinting and path-loss method.*

	Fingerprinting method	**Path-loss method**
Data storage needs	Very high, infeasible for large areas	Small to moderate even for large areas (only few path-loss parameters per AP/BS are stored, not the full measurement database)
Processing	Euclidian distance plus Nearest Neighbor approach	Euclidian distance plus Nearest Neighbor or Bayesian approaches
Accuracy	Typically higher than for path-loss method	Lower than for fingerprinting methods; sometimes similar performance, environment dependent

3.4. Advantages and Disadvantages in RSS-based location estimation

RSS-based location estimation has several advantages over the other localization methods [15]. The RSS-based location estimation can be implemented with the existing wireless networks such as WLAN and cellular networks. This does not require the deployment of dedicated hardware infrastructure solely for indoor and urban positioning where the GPS positioning often fail to offer a position estimate. So, it requires no extra cost for deploying the wireless networks. All that is needed is the ability to read the RSS on the receiver, which is now-a-days available in most of the wireless devices. The modulation method, data rate and system timing precision are not relevant. Synchronization between the transmitter and receiver is not required as in other TOA/TDOA principles [15].

On the other hand, RSS-based location estimation has several disadvantages also. Since those WLAN and cellular signals are largely affected by interference and multipath, location accuracy is generally less than what achieved using the TOA/TDOA principles [15]. If the location estimation using RSS database comparison is used, a huge specific database must be created for any given location, which is time consuming and often tedious [27] due to extensive survey measurements. The biggest drawback of RSS database system is that they are site/environment dependent. The RSS database is applicable only for the particular site where it has been taken. It requires a new database for new site/environment and even the physical or topological change in the same environment may require creating a new database [16], [27].

4. Channel measurements

From the previous chapters, we have seen that the RSS from the cellular and WLAN networks can be used as an alternative method to estimate user location where the GPS signals fail to serve. Compared to other techniques such as AOA or TOA, RSS-based estimation is more effective in indoors due to multipath and shadowing effect which generates the non line-of-sight (NLOS) environment that causes inaccurate angle and time estimation [19]. RSS can easily be measured with a device integrated with a wireless network interface card which has the ability to measure RF signals. So, RSS-based methods are the simplest ones to be used among all other variants.

Although, RSS-based method seems to be effective and simpler, more challenges arise due to fact that the RSS fluctuates significantly over time and space. The RSS is very much influenced by the environment because of fading, scattering and shadowing. So, the errors incurred through this method could be significantly varying too. Several characteristics of RSS have not been thoroughly investigated yet. Many researchers have relied on the RSS as sensor information to determine objects location while ignoring the characteristics of RSS. Therefore, understanding the actual RSS distribution could be a key to improve the performance of indoor positioning systems [19]. This chapter presents our studies regarding the RSS channel measurements, such as shadowing effect Ψ, path-loss exponent n, apparent transmit power P_T of APs or BSs for both indoor and outdoor WLAN and cellular networks.

4.1. Measurement setup

In order to analyze the channel parameters such as shadowing effect Ψ, path-loss exponent n, real-time measurements have been taken in different scenarios, for example indoors, outdoors, malls, university, parks etc. The author has also participated in the measurement campaigns by collecting some of the data used in this thesis. The Author also had access to a database of data collected by other colleagues in the research group.

For all the indoor WLAN measurements, the measurement data has been collected using a Windows tablet running a software tool that converts the map location into a Cartesian coordinates. The position on the map had been manually added during the data collection phase. For the outdoor WLAN and cellular cases, the measurements have been collected

using a Nokia mobile phone with an incorporated GPS receiver [4]. The map position was based on the GPS estimates. *Figure 4.1* below shows the fingerprint measurement data for indoor WLAN in one of the floor in the Tampere University of technology (TUT) building. Similarly, *Figure 4.2* shows the fingerprint measurement data for outdoor GSM case in a suburb in Tampere city (Hervanta district).

Figure 4.1 *Example of measured fingerprint data for indoor WLAN networks (2nd floor TUT).*

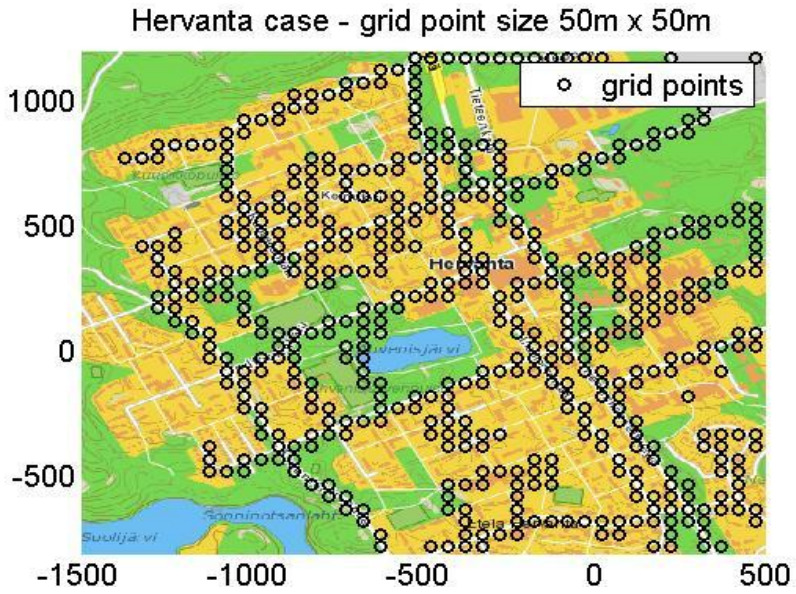

Figure 4.2 *Example of the measured fingerprint data for outdoor GSM networks.*

Once the data has been collected, it was also mapped to a discrete rectangular grid, of L x L dimensions (e.g., L=2-10 m in WLAN case and L=25-50 m in cellular case) for the purpose of data compactness. An example of synthetic grid mapping is showing in *Figure 4.3*.

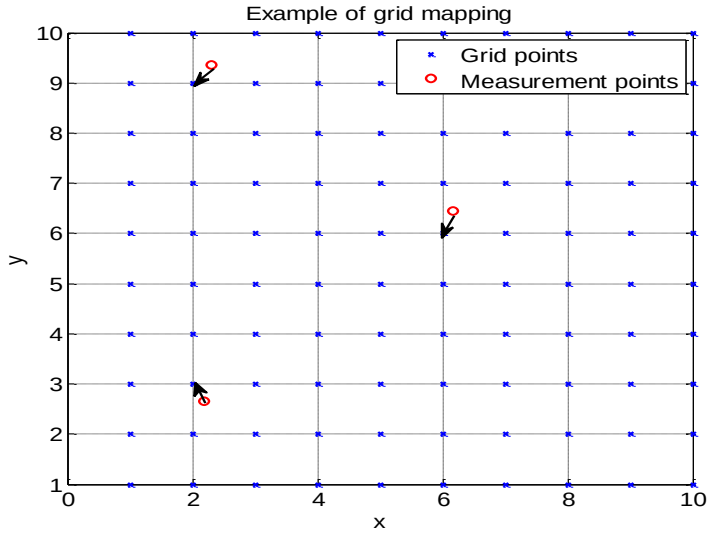

Figure 4.3 *Example of synthetic grid mapping for the fingerprint data.*

4.2. Shadowing modeling via fixed point approach

RSS varies over time due to small movements of MS or objects in the propagation path or other physical channel effects such as diffractions, refractions, reflections, etc [16]. If we can capture repeated RSS measurements at the same reference point from a series of transmissions from all available APs/BSs, then we can measure the fluctuations in the RSS. The fluctuation in the RSS gives the joint shadowing and fading effect in that radio channel. Thus, for this, we used a fixed point approach. This approach consisted in the fact that the measuring device, (tablet of Nokia mobile phone) was kept stationary, in a fixed point on a table, a desk, a shelf or a car chair, for a duration ranging from 1 to 10 hours and was capturing continuously, every 1s, both WLAN and cellular signals (either 2G or 3G). Assuming that the measurement conditions remain the same in fixed point measurements, the fluctuations in the RSS levels are solely given by the shadowing effect (or, more exactly, the cumulated shadowing-fading effects). Therefore, the shadowing part Ψ_{ap} (in dB) can be estimated as:

$$\Psi_{ap} = P_{ap} - E(P_{ap}), ap = 1,2, \dots N \tag{4.1}$$

where, P_{ap} (in dBm) is the measured RSS from AP or BS ap and $E(P_{ap})$ is the mean of all the scanned RSS from certain AP/BS ap . Thus, shadowing standard deviation is simply the standard deviation of the measured RSS minus its mean value i.e. $std(\Psi_{ap})$. This has been calculated for all the hearable APs/BSs respectively for WLAN/cellular case and the statistics were taken as averages over all APs or BSs available.

In what follows we show the results for indoor GSM (*Figure 4.4*), indoor WCDMA (*Figure 4.5*) and indoor WLAN (*Figure 4.6*) cases. Those plots show the mean RSS and shadowing standard deviation in indoors, cellular and WLAN cases. In the plots, each dot corresponds to an AP in WLAN and BS in cellular case in a scenario. A larger numbers of APs were heard in the WLAN case than the number of BSs in cellular case, which explains the dense plot in WLAN case in *Figure 4.6*. The x-axis gives the measure of the standard deviation of shadowing in dB while the y-axis gives the mean RSS in dBm of that AP/BS represented by that dot.

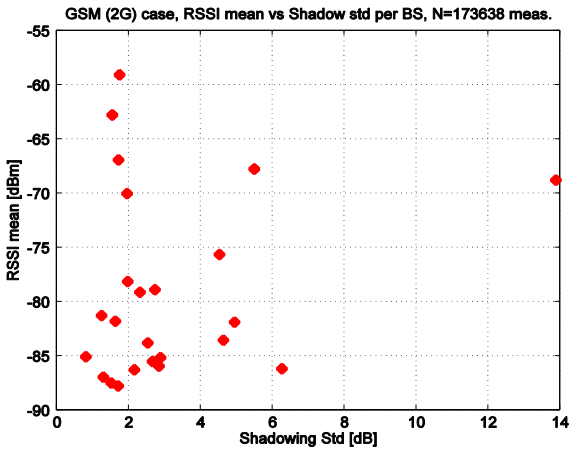

Figure 4.4 *Mean RSS versus standard deviation of RSS for indoor GSM case.*

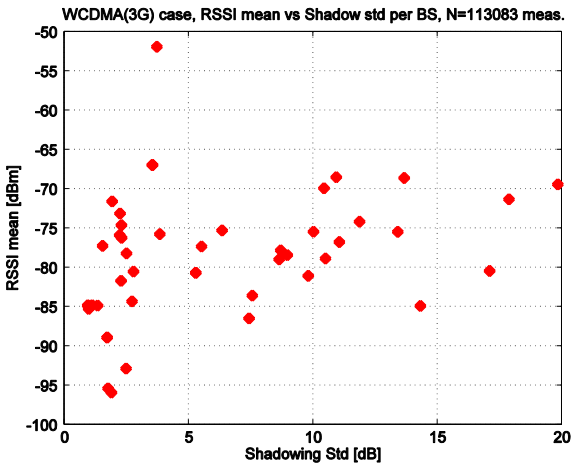

Figure 4.5 *Mean RSS versus standard deviation of RSS for indoor WCDMA case.*

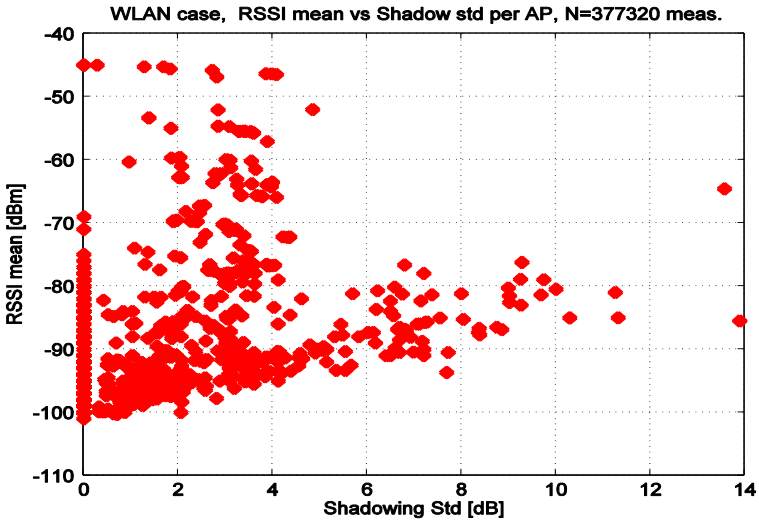

Figure 4.6 *Mean RSS versus standard deviation for indoor WLAN case.*

From the observations, the shadowing standard deviation in GSM case was around 6dB, for WLAN case it was seen to be somewhere up to 12dB and for WCDMA even up to 20dB. This has shown that the shadowing effect for WLAN signal is more than for GSM and is largest for the WCDMA case.

We have also noticed that there is no relationship between the mean RSS and the shadowing standard deviation, meaning that both distant and close measurements are likely to be affected by the same shadowing effects.

The outdoor cases for WCDMA and WLAN cases are illustrated in *Figure 4.7* and *Figure4.8*. Both WLAN and WCDMA exhibit quite small shadowing deviations of approximately 2.5 – 3dB outdoors.

This may be explained by less closely-spaced obstacles in outdoor environments compared to indoor environments and therefore scatters less and less RSS fluctuations.

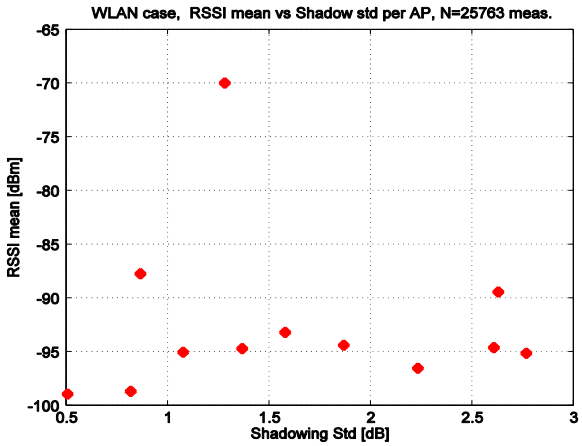

Figure 4.7 *Mean RSS versus standard deviation for outdoor WLAN case.*

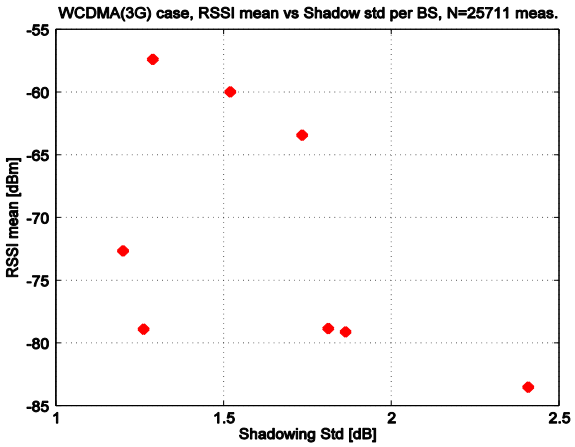

Figure4.8 *Mean RSS versus standard deviation for outdoor WCDMA case.*

From the theory, it is known that the shadowing effect is typically modeled using Gaussian (or normal) distribution in logarithmic scale with zero mean and certain standard deviation. An example from one of the measurements for indoor WCDMA the PDF plot for shadow fading is as shown in the *Figure 4.9* below and its measured probability distribution function (PDF) is compared with a theoretical Gaussian PDF of zero mean and same standard deviation as the standard deviation of the measured data. We can see that the Gaussian curve is a good fit, as predicted by the typical theoretical models [11]. From our data, it has been observed that the shadowing distribution follows the Gaussian distribution in most of the cases.

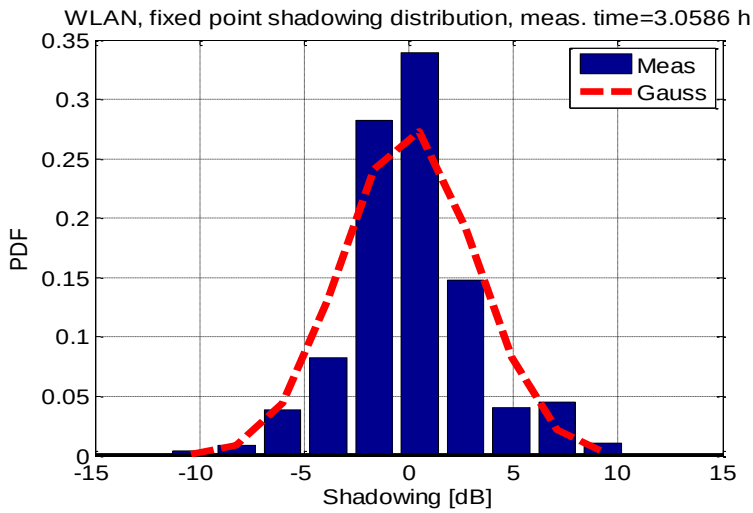

Figure 4.9 *Example of shadowing PDF (measured and theoretical fit), good Gaussian fit case.*

However, there were also several exceptions noticed, as illustrated in *Figure 4.10*. Here, also the theoretical Weibull PDF is shown, because Weibull has been also used in some papers to characterize the shadowing [22]. From our observations, when Gaussian distribution failed to match the measured data, neither the Weibull distribution was offering a better match. However, these measured cases where Gaussian curve was not giving a good fit, were rare and, in our simulator, we only incorporated the Gaussian shadowing at the moment.

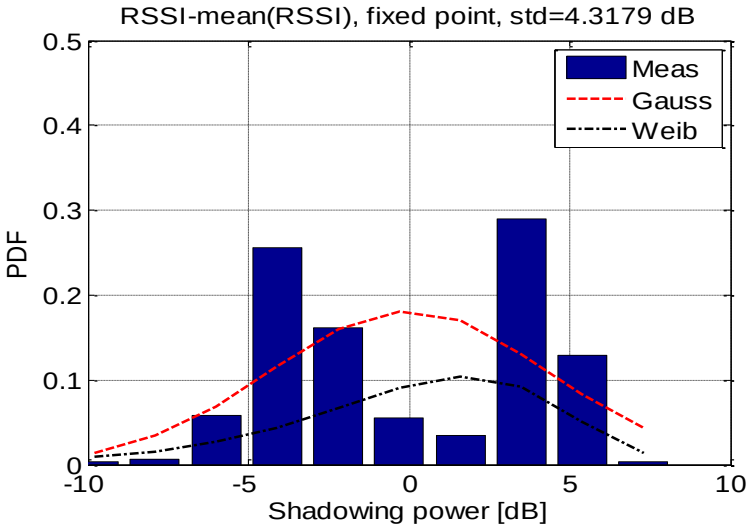

Figure 4.10 *Example of shadowing PDF (measured and theoretical fit), bad Gaussian fit case.*

4.3. Path-loss parameters estimations

In the training phase we simply collect the RSS database by doing several measurements around certain, desired, geographical area. In such environments, the APs/BSs locations and their transmit power are not known beforehand. So, in this section, the estimation of APs/BSs location and their path-loss parameters have been investigated, both for indoor and outdoor WLAN and cellular networks. The simplest path-loss model as shown in equation (4.2) [4], [23], that gives the RSS at the receiver in terms of the path-loss coefficient, transmit power and the location of AP/BS.

$$P_{i,ap} = P_{T,ap} - 10nlog_{10}(d_{i,ap}) + \Psi_i \tag{4.2}$$

where, $P_{i,ap}$ is the received power in dBm at the i^{th} measurement point, $P_{T,ap}$ is the transmitted power in dBm from the ap^{th} AP, n is the path-loss exponent, Ψ_i is the shadowing in dB and the distance between the ap^{th} AP and the i^{th} measurement point is:

- in 2D-model case:

$$d_{i,ap} = \sqrt{(x_i - x_{ap})^2 + (y_i - y_{ap})^2} \tag{4.3}$$

  where, (x_i, y_i) is the 2D coordinate of the i^{th} measurement point and (x_{ap}, y_{ap}) is the coordinate of the ap^{th} AP.

- in 3D-model case:

$$d_{i,ap} = \sqrt{(x_i - x_{ap})^2 + (y_i - y_{ap})^2 + (z_i - z_{ap})^2} \tag{4.4}$$

  where, z_i and z_{ap} are the height coordinate of the i^{th} measurement point and the ap^{th} AP respectively.

for i =1,….N measurement points index and ap =1,…N_{ap} AP index.

Since the height dimension was not accurate neither from the tablet maps nor from the GPS coordinates 2D model was used for the measured data.

Firstly, from the collected measurements, the APs/BSs location was estimated by taking an average over the positions of the measurements with the highest RSS. Then also the path-loss coefficient and the apparent transmit power of those APs/BSs were estimated using

some form of regression, such as the linear regression, as illustrated in the 2 figures below (*Figure 4.11* and *Figure 4.12*).

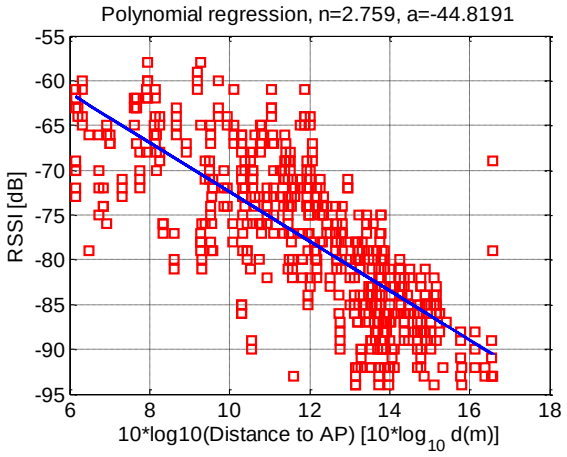

Figure 4.11 *Example1 of linear regression for estimating path-loss exponent and transmit power.*

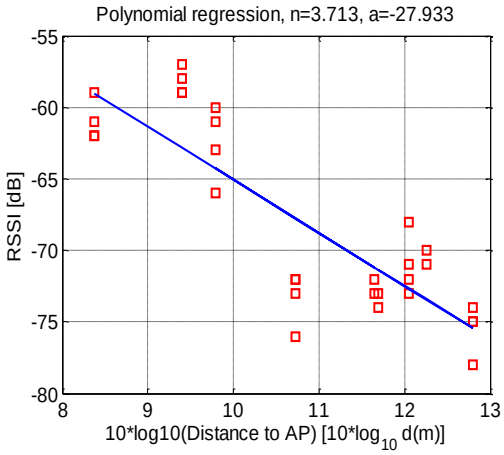

Figure 4.12 *Example2 of linear regression for estimating path-loss exponent and transmit power.*

The results of the estimated parameters based on real-time channels are presented below in the tables. The mean, standard deviation and maximum of estimated path-loss coefficient *n* and mean and standard deviation of estimated apparent transmitted power P_T and the total number of heard APs/BSs per scenario are presented. In *Table 4.1*the results for indoor WLAN are presented. Similarly, in *Table 4.2* the results for outdoor GSM and in *Table 4.3* the results for outdoor WCDMA are presented respectively.

Here, in the tables, a district means a part of town different from downtown. The statistics have been done on a much smaller number of APs/BSs than the total number of heard APs/BSs because, only those APs/BSs that are heard in at least 10 different measurement points and for which we obtained positive estimation of n have been kept, in order to have more reliable results.

Table 4.1 *Estimated path-loss parameters for indoor WLAN networks.*

	Mean $\tilde{n}$	**Standard deviation** $\tilde{n}$	**Max** $\tilde{n}$	**Mean** $\tilde{P}_{T,ap}$ **[dBm]**	**Standard deviation** $\tilde{P}_{T,ap}$ **[dB]**	**Total number of heard APs**
Univ. building 1, 1st floor	2.49	1.89	10.3	-50.13	25.36	190
Univ. building 1, 2nd floor	2.50	2.02	11.8	-49.73	29.84	205
Univ. building 2, 1st floor	1.38	1.02	4.5	-62.88	19.43	244
Univ. building 2, 2nd floor	1.43	1.08	5.4	-64.14	18.89	331
Tampere shopping center 1, 1st floor	1.61	1.03	4.2	-59.01	23.25	64
Tampere shopping center2,1st floor	1.94	1.28	4.0	-52.73	25.57	31
Tampere Mall, 2nd floor	1.42	1.27	4.7	-66.35	24.20	158
Tampere Mall, 3rd floor	1.25	1.42	7.6	-69.63	24.79	188

An example of the probability distribution function of the estimated path-loss coefficient is illustrated in *Figure 4.13* (below) for TUT building.

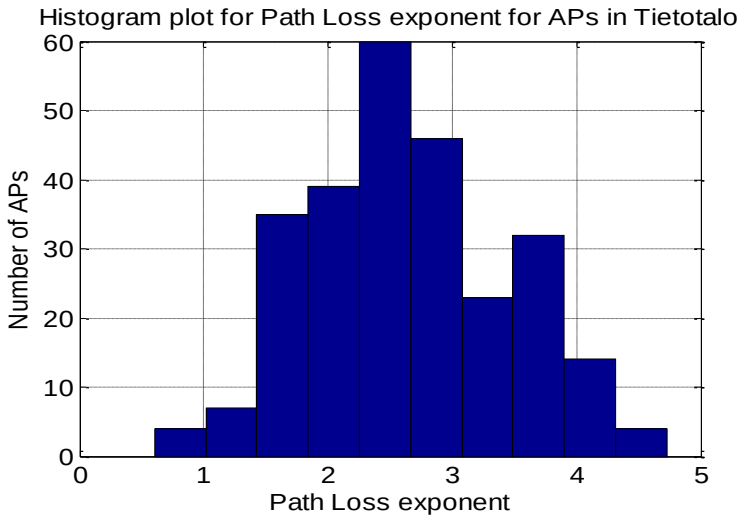

Figure 4.13 *PDF of estimated path-loss coefficient inside one of the TUT building.*

Table 4.2 *Estimated path-loss parameters for outdoor GSM networks.*

	Mean $\tilde{n}$	**Standard deviation** $\tilde{n}$	**Max** $\tilde{n}$	**Mean** $\tilde{P}_{T,ap}$ **[dBm]**	**Standard deviation** $\tilde{P}_{T,ap}$ **[dB]**	**Total number of heard BSs**
Tampere large district	1.73	0.85	4.82	-28.81	23.37	47
Bucharest park	0.79	0.81	3.26	-65.05	19.26	21
Bucharest downtown	2.19	1.91	8.71	-19.04	51.17	77
Bucharest large district	1.73	1.51	10.08	-26.67	49.76	81

Table 4.3 *Estimated path-loss parameters for outdoor WCDMA networks.*

	Mean $\tilde{n}$	**Standard deviation** $\tilde{n}$	**Max** $\tilde{n}$	**Mean** $\tilde{P}_{T,ap}$ **[dBm]**	**Standard deviation** $\tilde{P}_{T,ap}$ **[dB]**	**Total number of heard BSs**
Rovaniemi downtown	1.25	0.76	2.98	-51.76	26.56	29
Munich small district	1.03	1.11	3.14	-60.83	33.18	7
Tampere large district	1.37	0.76	3.88	-46.73	22.77	77

From the observations, for outdoor cellular cases, the mean estimated path-loss coefficient n is rather small, around 1-2. The mean estimated apparent transmit power for GSM is a little bit higher than for WCDMA. For indoor WLAN, the average estimated path-loss coefficient is approximately 1.5 - 2.5 and the mean apparent transmit power P_T is around -50 dBm to -70 dBm.

Hence, this shows that the WLAN signals, especially indoor WLAN signals, suffer of slightly higher attenuation than the cellular ones. Surprisingly also, the average path coefficient that best fits our measured models was typically beyond the 2-factor of the free-space path-loss model. This might be explained by the fact that our simplified path-loss model does not take into account the environment geometry (i.e., obstacles, buildings, floors, walls).

4.4. Shadowing models via path-loss model

From the principle of wireless communication and signal propagation, an RF signal attenuates over a distance as it propagates through the space. The RSS at a receiver can be predicted by the path-loss model given by equation (4.2) and (4.3) in the above section.

For this method, from the collected location fingerprints, the APs/BSs location and the apparent transmit power P_T and path-loss coefficient n were estimated. Once those path-loss parameters have been estimated then we laid the path-loss grid of uniform grid size and recalculated the RSS values at those grid points using the same path-loss model expressed above in equation (4.2) but now without any shadowing. Now when the recreated path-loss

radio map was subtracted from the radio map from measured RSS, the shadowing effect was obtained. Here, the measured RSSs minus the RSSs due to recreated path-loss model forms the shadowing effect. The following *Figure 4.14* gives a clear illustration

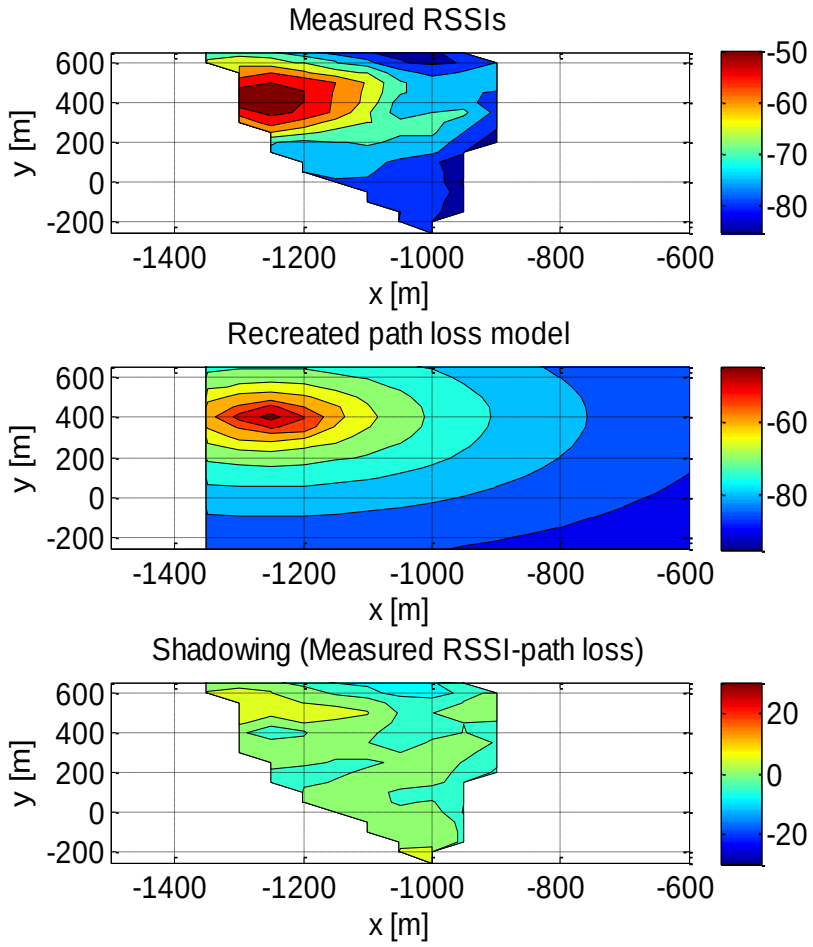

Figure 4.14 *Steps involved in shadowing modeling via path-loss method.*

In *Figure 4.14 above*, the first sub plot shows the radio map of true measured RSS, the second subplot shows the recreated path-loss radio map after estimating AP location, *n* and P_T and the third subplot gives the shadowing effect after subtracting the recreated path-loss radio map from the measured radio map.

Using this approach, the shadowing standard deviation was calculated for all those APs/BSs heard in at least 10 measurement points at that environment. The following *Figure 4.15* and *Figure 4.16* below show the shadowing standard deviation for different GSM outdoors, WCDMA outdoors, WLAN indoors and outdoors scenarios for a wide range of measurements done in four countries (Finland, Germany, Italy and Romania) and 5 different towns (Tampere, Rovaniemi, Munich, Sorrento and Bucharest). The standard

deviation of shadowing presented in the bar diagram below is the average over shadowing standard deviation of all the APs/BSs considered in the calculation.

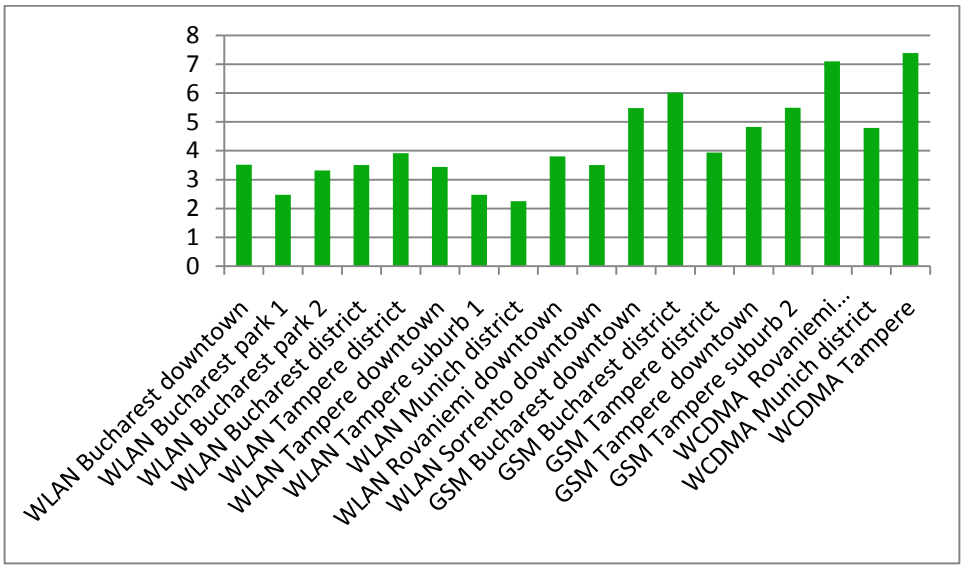

Figure 4.15 *Graphical representation of shadowing standard deviation for outdoor cellular (2G and 3G) and WLAN cases.*

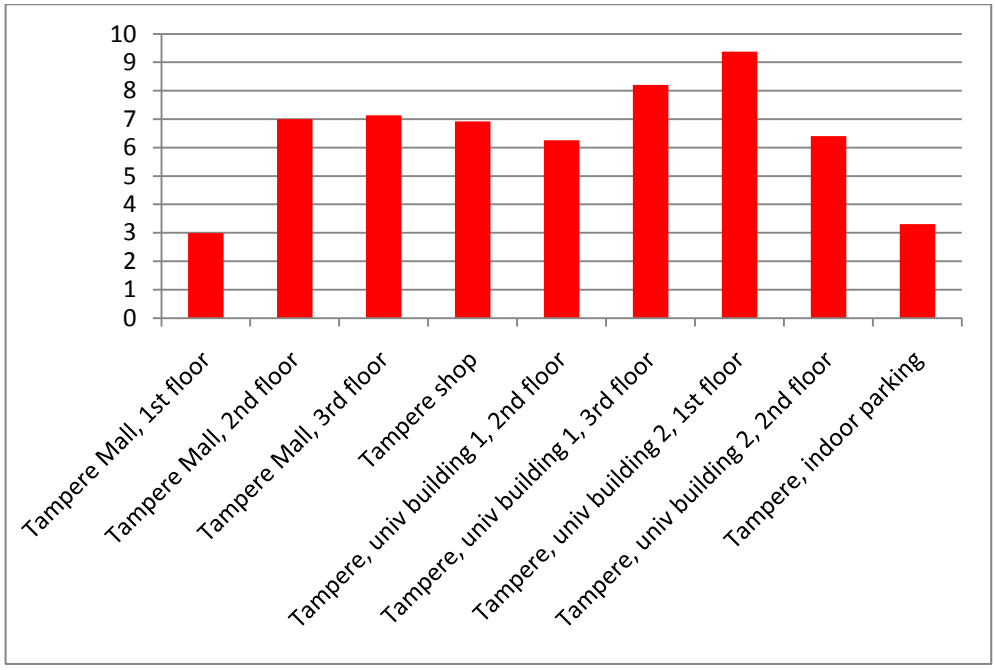

Figure 4.16 *Graphical representation of shadowing standard deviation for indoor WLAN cases.*

Figure 4.15 shows the shadowing standard deviation in the outdoor cases for WLAN, GSM and WCDMA scenarios. From the plots, the shadowing standard deviation for outdoor

WLAN is around 2.5 - 4 dB, while for GSM outdoor seems to be around 4 - 6 dB and for WCDMA at around 5 - 7.5 dB (i.e., higher fluctuations of the WLAN signal compared to cellular signal). *Figure 4.16* shows the shadowing standard deviation in the indoor WLAN cases. In the indoor WLAN cases, the shadowing varies from 3 - 9 dB. From this observation also, it is seen that the shadowing effect is high in the indoor environment than in outdoor cases. The fluctuations of cellular signal indoors was not measured because the tablet used for indoor measurements was equipped only with WLAN receiver (not also with cellular receiver) and the mobile phones had no map reference for the coordinates indoors.

4.5. Decorrelation distance in correlated shadowing

The shadowing is caused by the presence of obstacles in the propagation path of the radio signals. The shadowing greatly depends on the terrain profile. It is quite true that the morphology structure and terrain profile do not change abruptly, so it is expected that the shadowing distribution would show some spatial correlation. Moreover, another user should experience similar shadow fading effects passing by the same location [13]. The distance up to which the spatial correlation seems to exist varies depending on the environment.

In this section we investigate the decorrelation distance in various scenarios: outdoor GSM and WLAN and indoor WLAN cases. An example of how the decorrelating distance is computed is given in *Figure 4.17*. Here the 3D dependence of RSS measured at different x and y distances is shown. The decorrelation distance is taken as the distance at which the autocorrelation decreases to half of its maximum value.

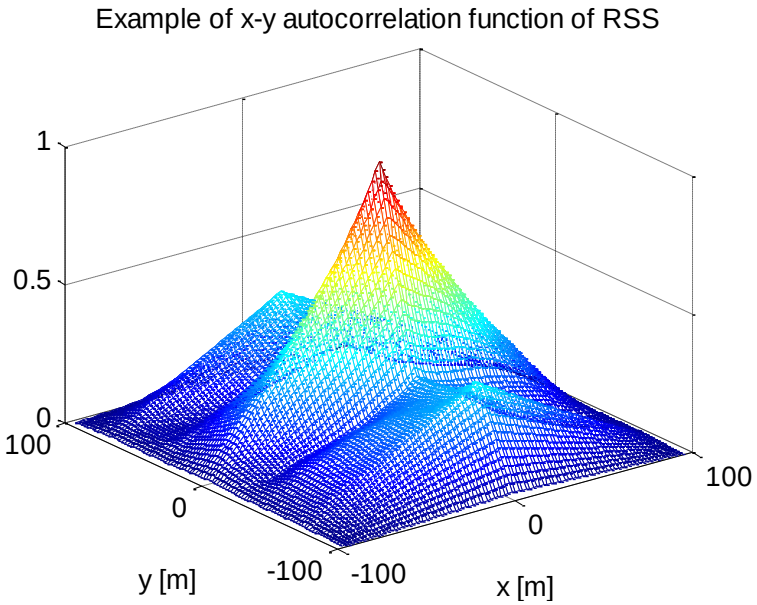

Figure 4.17 *Example of correlation matrix of RSS.*

The decorrelation distances on x and y axes were first computed, as illustrated in *Figure 4.18*, and the total decorrelation distance was taken as the average over the two.

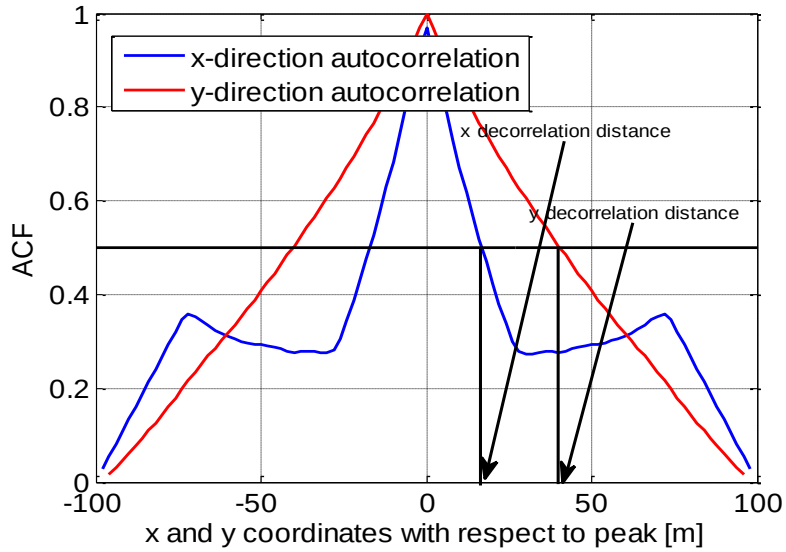

Figure 4.18 *x-y Decorrelation distances illustration.*

Figure 4.19 shows the decorrelation distance observed in the outdoor GSM cases and *Figure 4.20* shows for the outdoor WLAN cases. The measured data used here are from Tampere, Finland and Bucharest, Romania. Similarly, *Figure 4.21* gives the observation for indoor WLAN networks in the buildings in Tampere.

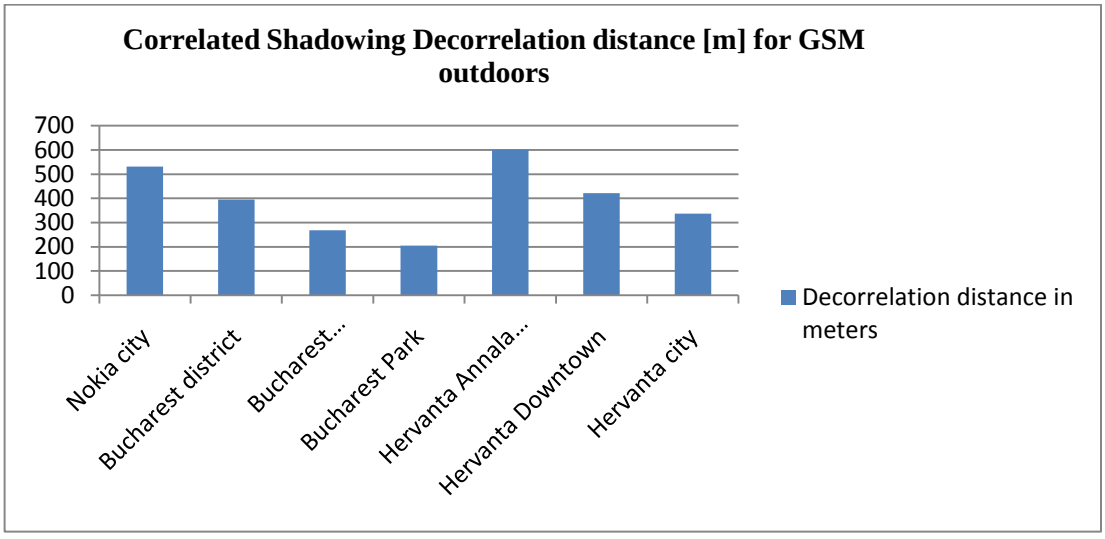

Figure 4.19 *Average decorrelation distance in meters for outdoor GSM cases.*

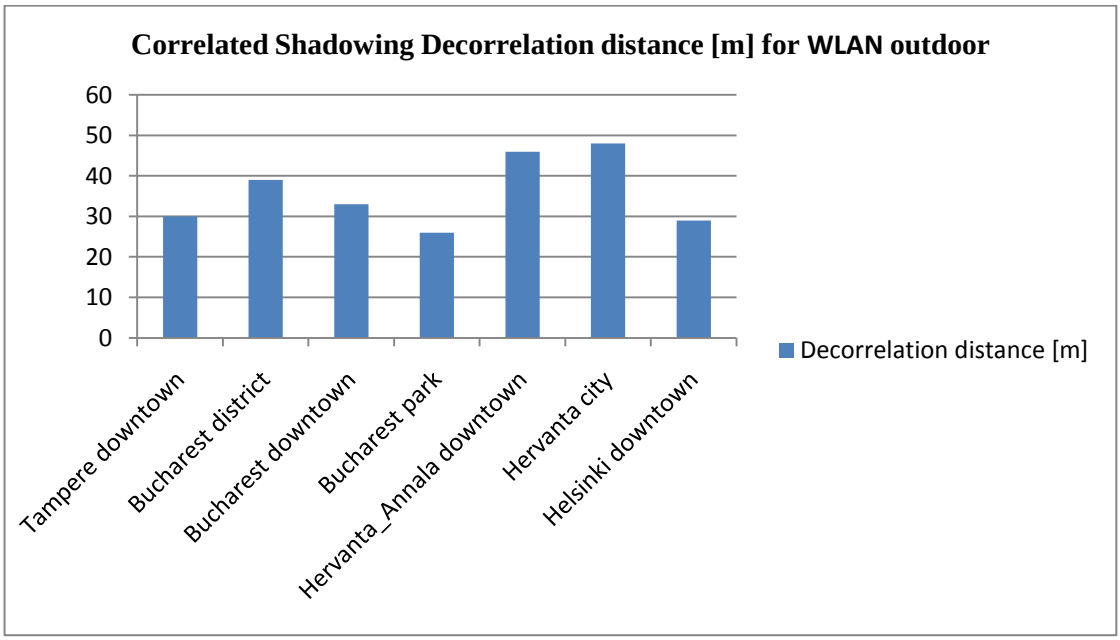

Figure 4.20 *Average decorrelation distance in meters for outdoor WLAN cases.*

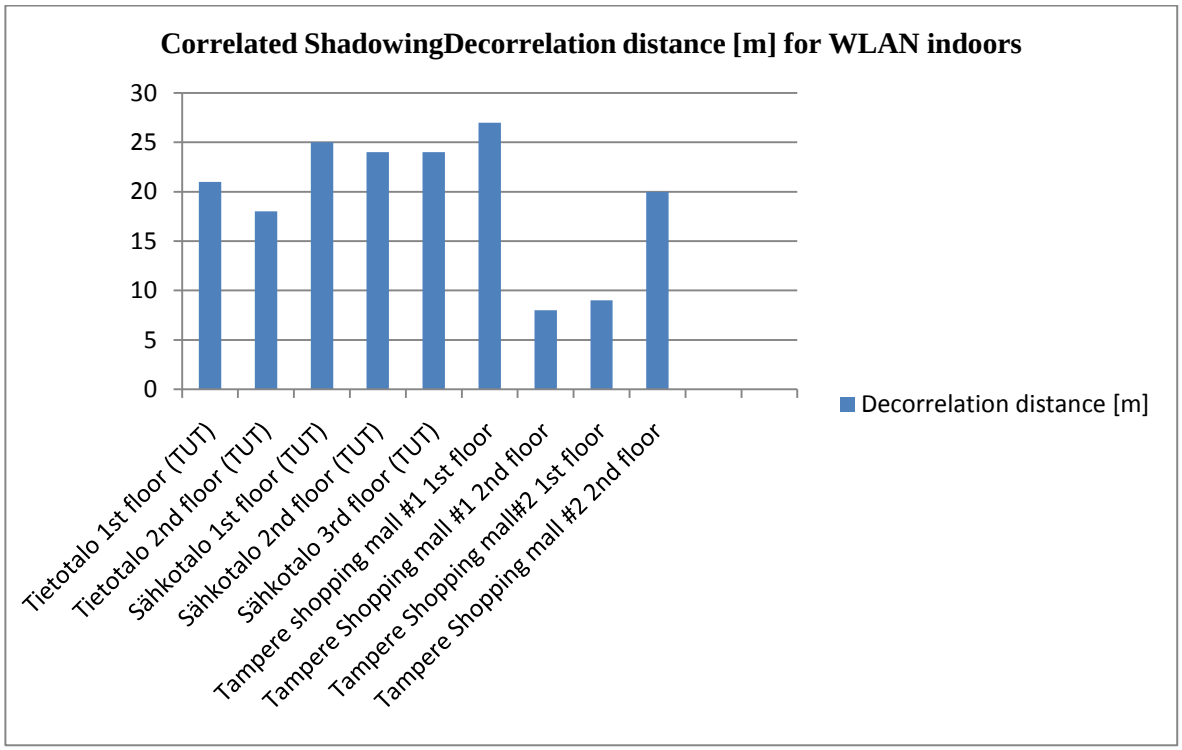

Figure 4.21 *Average decorrelation distance in meters for indoor WLAN cases.*

From the charts, it is seen that the decorrelation distance for the cellular (GSM) networks is the highest compared with WLAN data. This means that the signal fluctuations tend to keep similar trends over wide distance ranges in GSM. This can be explained by the fact that the cellular networks have larger coverage than compared to the WLAN networks. For the GSM outdoor cases, the decorrelation distance seems to vary from 200 - 600m. And for the outdoor WLAN, it ranges from 25 -50m while for indoor WLANs it can range from 8 - 25m. The estimations are roughly based on the observations from the measured data. WLANs exhibit much smaller decorrelation distance as compared to cellular networks. In the outdoor scenarios, smaller average decorrelation distance is observed in urban/dense buildings areas than in open spaces/less dense buildings areas, as expected. Knowledge about shadowing decorrelating values can for example help in better modeling the path-losses needed for RSS-based mobile positioning.

4.6. Power maps and scatter diagrams based on measured data

4.6.1. 2D models

Here, we have presented some power maps from the measured data in 2D case. *Figure 4.22* shows the power map from a BS in outdoor GSM case, *Figure 4.23* shows the power map for a WCDMA BS in Tampere city. Similarly, power map from an AP in indoor WLAN is presented in *Figure 4.24* and from an AP in outdoor WLAN for Tampere city is shown in *Figure 4.25*. The 2G GSM case and 3G WCDMA depict similar kind of power maps with RSS ranging from -50 to -110 dBm. while for the WLAN case the RSS ranges from -20 to -100dBm indoor and -70 to -90 dBm outdoor. The RSS seems to be quite weak in the outdoor WLAN case. It also largely attenuates over a shorter distance moved away from the AP both in indoor and outdoor cases. This also, shows a higher shadowing deviation in the indoor cases.

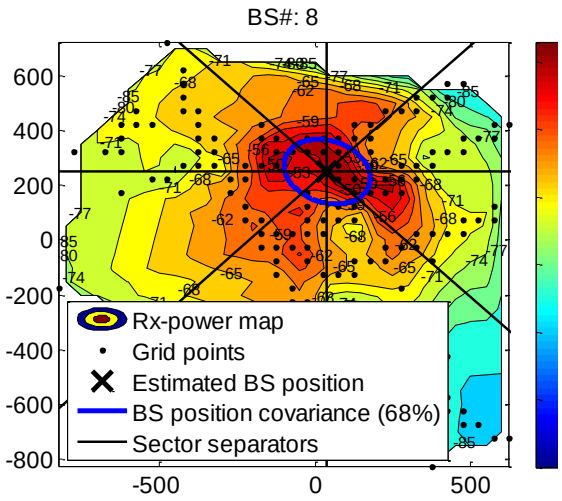

Figure 4.22 *Example power map from a BS in outdoor GSM case (Hervanta district).*

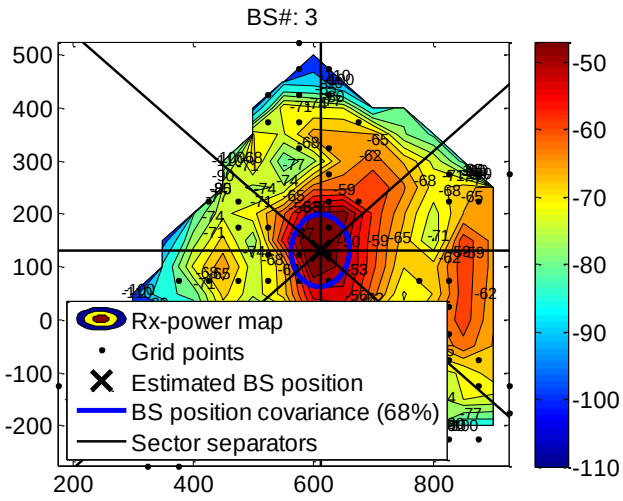

Figure 4.23 *Example power map from a BS in outdoor WCDMA case (Tampere city).*

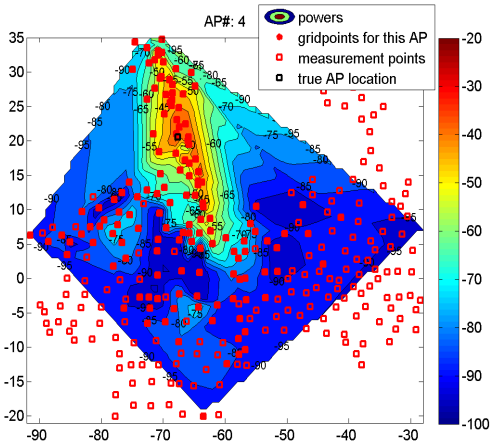

Figure 4.24 *Example power map from an AP in indoor WLAN case (inside TUT building).*

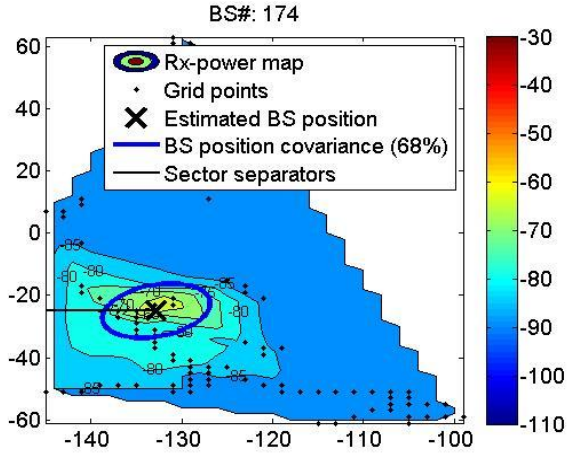

Figure 4.25 *Example of power map from an AP in outdoor WLAN case (Tampere city).*

4.6.2. 3D models

Here, the 3D models for multi-floor buildings were investigated. An example of scatter plots for such multi-floor building is shown in *Figure 4.26* for a 4-floor university building in Tampere, Finland.

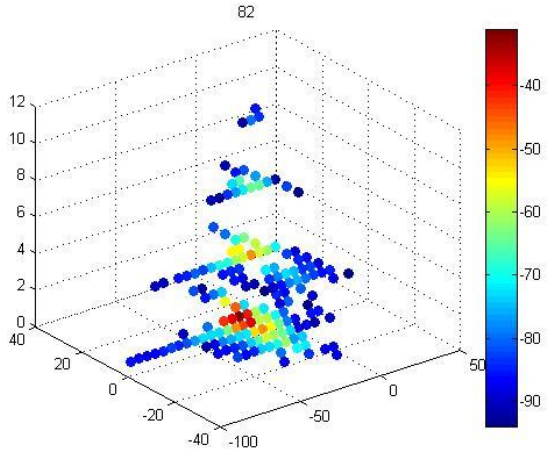

Figure 4.26 *Example of power scatter diagram for indoor WLAN case (multi floor university building without open spaces).*

Another example of scatter plots for a multi-floor building with open spaces and 6 floors (a mall in Finland) is shown in *Figure 4.27*.

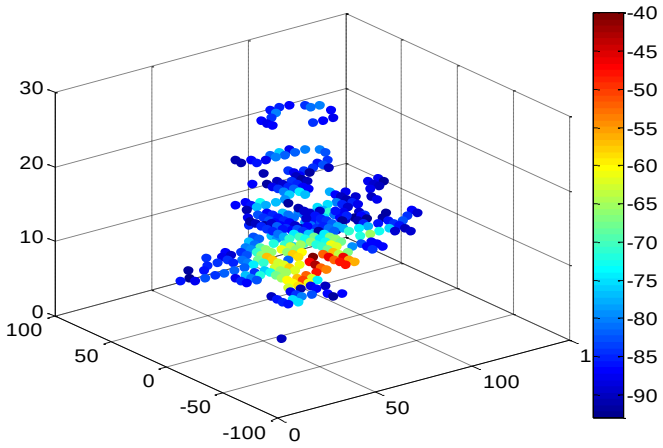

Figure 4.27 *Example of power scatter diagram for indoor WLAN case (multi floor mall building with open spaces).*

Typically, the range of WLANs is much lower than for the cellular case. In the above example *Figure 4.26* the same AP has been heard over all 4 floors of the building, but at lower power levels than on the floor where AP was situated. In the second example *Figure 4.27* the same AP has been heard across all 6 floors, due to the open spaces inside the mall where the measurements have been done.

5. MATLAB simulator

It has been seen from Chapter 3 and Chapter 4 and also from literatures [16], [26] that the RSS-based location estimation with measurement data often requires large measurement database for statistical significant results. Keeping in view this shortcoming of the measurement-based analysis and location estimation, a simulator in MATLAB has been developed. This MATLAB simulator not only generates various RSS databases (with adjustable parameters, chosen in our model according to the measured data), but it also implements the position estimation using both the fingerprinting method and the path-loss method. The positioning error/accuracy is then expressed in RMSE based on the distance error. This chapter gives a brief overview of the MATLAB simulator in terms of how the simulator has been developed and what are its objectives and shortcomings. The MATLAB simulator is the main contribution of this thesis. The simulator has been mainly built by the Author, with some input from the supervisor and the other team members.

5.1. General Description of the MATLAB simulator

There are two phases in the RSS-based location estimation a) training phase and b) estimation phase. In the same way, there are two phases in the MATLAB simulation also:

- Data generation phase,
- Estimation phase

In real-time/measurement-based positioning, the training phase is carried out with the RSS database collection in the experimental site. But in the MATLAB simulator, firstly we generate (randomly) the measurement points and the building structure (e.g., number of APs per floor and per building, number of APs with multiple Service Set Identification SSID, etc) and then we compute the RSS value using the simplified omni directional path-loss model as given in equation (4.2), (4.3) and (4.4) from section 4.3, Chapter 4.

In accordance to the measurement data analysis, correlated Gaussian shadowing Ψ_i is added in case of 2D model. Because of the difficulties in building 3D filter to filter the white noise model of shadowing, in the 3D case only the uncorrelated Gaussian shadowing is currently modeled. In both cases, the shadowing can have variable standard deviation (e.g. 4-12 dB based on measurement results).

In our model, also in accordance with the measurement data analysis from Chapter 4, we also introduced a certain probability to hear an AP at a certain distance, according to the following rule:

$$p = \begin{cases} 1 & if\ d_{i,ap} \leq d_{min} \\ p_{interp} & if\ d_{min} < d_{i,ap} \leq d_{max} \\ 0 & if\ d_{i,ap} > d_{max} \end{cases} \quad (5.1)$$

where $d_{i,ap}$ is the distance between the ap^{th} AP and the i^{th} measurement coordinate calculated as given by equation (4.3) in 2D model and (4.4) in 3D model, d_{min} and d_{max} are some minimum and maximum distance bounds defining the hearability limit of an AP (currently set at 12 m and 40 m, respectively, in accordance with some of the measurement data from Chapter 4) and p_{interp} is the interpolated probability, which gives us the probability to hear an AP if its distance to the measurement point is in between the minimum and maximum limits. If $d_{i,ap}$ is less than the minimum distance d_{min} then the AP is assumed to be always heard, while if $d_{i,ap}$ is higher than the maximum distance d_{max}, then the AP is never heard. In between these distances, it is heard with a certain non-zero and non-unit probability. This modeling is needed in order to take into account various building effects, such as walls and doors within a building. The threshold for hearability of an AP in real indoor WLAN measurements has been presented in *Figure 5.1* below.

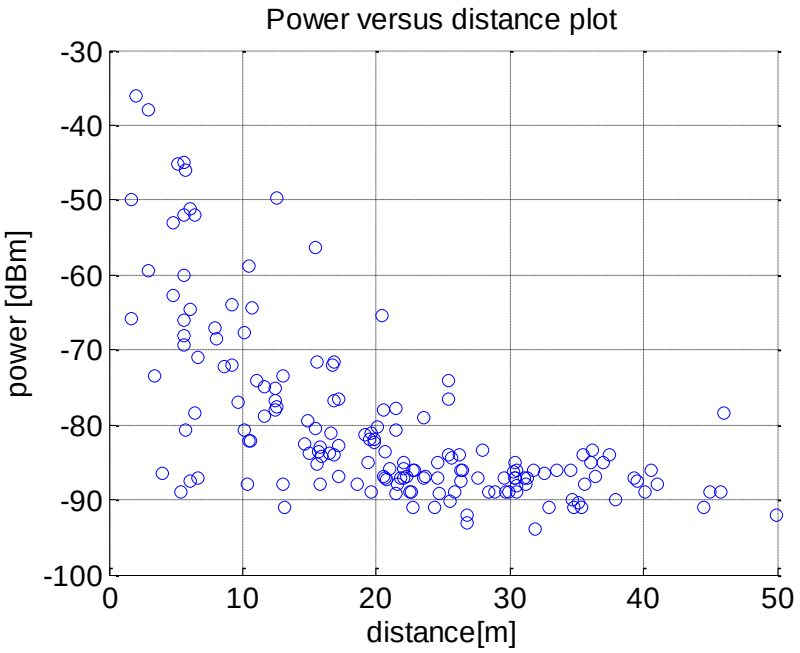

Figure 5.1 *Example plot of power versus distance from an AP in real indoor WLAN case.*

Hence, in this way the RSS database is created.

In order to generate the fingerprint data, firstly the adjustable parameters like number of fingerprints, length, breadth and height of the area, number of user points, number of APs, shadowing standard deviation, decorrelation distance are fed to the simulator. Then accordingly, it will create the fingerprint database. *Figure 5.2*, *Figure 5.3*, *Figure 5.4* and *Figure 5.5* below show how the fingerprint data looks like:

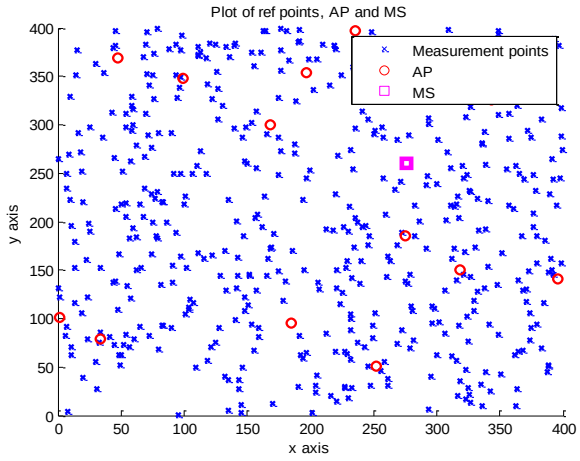

Figure 5.2 *Randomly located fingerprint data, APs and a MS in 2D plane.*

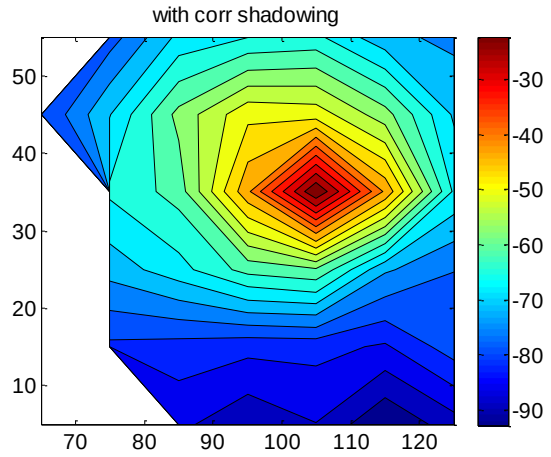

Figure 5.3 *Fingerprints in terms of Radio map due to a single AP in 2D plane.*

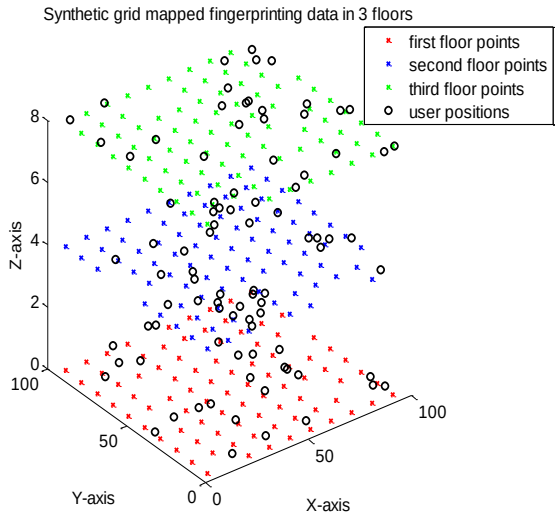

Figure 5.4 *Fingerprint data and user points in 3 floors in 3D case.*

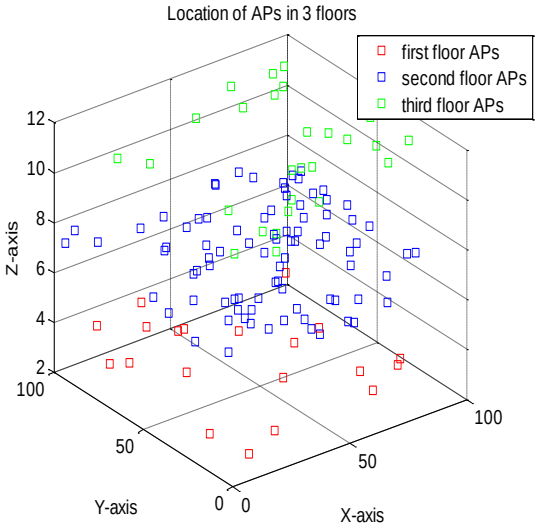

Figure 5.5 *Location of APs in 3 floors in 3D case.*

In *Figure 5.2* the randomly located fingerprints in 2D in 400m x 400m square area has been shown. *Figure 5.3* shows the RSS in terms of radio (power) map from one of the AP in the fingerprint data. In *Figure 5.4*, similar kind of location fingerprints in 3D per square floor of 100m x 100m also mapped to a 10 x 10m synthetic grid has been shown. And *Figure 5.5*

presents a snapshot of randomly located APs per floor in the three- floor case presented in *Figure 5.4*. These figures show some of the generated fingerprint data in 2D and 3D case. These are only some examples for a random scenario. For simulations purposes, hundreds of such random scenarios have been generated and created.

In addition, for 3D multi-floor building, a floor attenuation factor has been added in the model. This is another random variable of the model and it is scenario-specific (e.g., equal floor attenuation per building):

$$P_{i,ap} = P_{T,ap} - 10nlog_{10}(d_{i,ap}) + \Psi_i - L_{floor} \quad (5.2)$$

where, L_{floor} is floor attenuation factor in dB e.g. between 5 and 35 dB added to above path-loss equation (4.2).

Then, once we have the RSS database, the estimation phase is similar as in the real time positioning. For the estimation phase purpose, additional data tracks with new measurement points (different from the training points) are generated, according to the same model as above. The shadowing standard deviation can also be different in the data tracks compared with the training data.

In this model, we have included the fingerprinting method and the path-loss model. In the fingerprinting method, Euclidian distance approach with the averaging over 5 nearest neighbors (NN) has been used in most of the cases. Some other metrics apart from Euclidian distance for finding the nearest neighbors in terms of RSS have also been investigated and they are to be presented in the next chapter.

For the path-loss model, firstly the location of APs, the apparent transmit power P_T and the path-loss exponents *n* are estimated using the approach described in chapter 4 section 4.3.

Once those path-loss parameters have been estimated, then we can create a the path-loss grid of uniform grid size and recalculate the RSS values at those grid points using the same path-loss model expressed above in equation (4.2) but now without any shadowing. Then the actual positioning using minimum Euclidian distance approach with 5 NN as in fingerprinting method has been used. This means that the training points are sorted in descending order according to the difference between the training RSS (i.e., stored in the database for fingerprinting or recalculated based on the path-loss parameters for the path-loss model) and the measured RSS. Then, we take the estimated MS position as the average over the coordinates of the 5 strongest training points (according to Euclidian distance criterion). It is to be remarked that also other distance criteria have been investigated, as shown in Chapter 6.

The positioning error is then expressed in terms of RMSE based on distance. The RMSE based on distance error is given as:

$$RMSE = \sqrt{\frac{1}{N_{points}} \sum_{i=1}^{N_{points}} (\Delta d_i)^2} \tag{5.3}$$

where,

- in 2D model case,

$$\Delta d_i = \sqrt{(x_{true} - x_{est})_i^2 + (y_{true} - y_{est})_i^2} \tag{5.4}$$

- in 3D model case,

$$\Delta d_i = \sqrt{(x_{true} - x_{est})_i^2 + (y_{true} - y_{est})_i^2 + (z_{true} - z_{est})_i^2} \tag{5.5}$$

for $i = 1,\ldots N_{points}$ number of user points, $(x_{true}, y_{true}, z_{true})$ is the true user position and $(x_{est}, y_{est}, z_{est})$ is the estimated user position.

5.2. Objectives of the MATLAB model

Therefore, to summarize our attained objectives with the created MATLAB model, we have:

- This simulator generates a real-time RSS-based positioning scenarios, which can be used to approximate and study the accuracy in the user positioning.
- The simulator provides simulated fingerprint data, which is modeled according to the realistic radio channel. This can avoid the extensive measurement campaigns, at least up to some extent.
- It can be used as a tool to do extensive research in RSS-based positioning, which is often hard to do with real measurements. The real measurements often lack all the information and also they are very time consuming to collect.
- Since the MATLAB simulator can also estimate the locations of APs, it can be used as a tool to estimate an appropriate location when we need to deploy a new AP in an already existing WLAN system in any indoor environment. Moreover, the geometry of AP (e.g., AP density per building and their relative location to each other) can be further analyzed in conjunction with the location accuracy.

5.3. Current shortcomings with the MATLAB model

Some of the shortcomings with the MATALB simulator are as follows:

- The model currently simulates the survey area as an open rectangular or square area based on the adjustable length and breadth parameter. It could not simulate the true indoor environment in terms of geometry, location of rooms, walls or furniture etc. The attenuations due to walls, furniture and other objects are modeled simplistically via two parameters:
    1. The floor attenuation parameter (randomly from one building to another)
    2. The hearable distance limit concept described in section 5.1.
- The location of fingerprints, APs and the user points all are randomly distributed, typically obeying a uniform distribution. This is still to be more thoroughly checked with more real time scenarios (since the available database with indoor measurements where maps were available has been quite limited).
- The current multiple service set identifier (SSID) support is limited to maximum 5 Media Access Control (MAC) addresses coming from the same location. This was also the maximum number observed from realistic measurement data done with mobile phones, but it may be still limiting more generic situations. However, this is a parameter easily adjustable in our current model.
- Even though the parameters such as the apparent transmit power P_T, the path-loss exponent n and the shadowing standard deviation are adjustable, the simulations have been carried out using the parameters based on the measurements in Chapter 4. Hence, those parameters may not be the universal parameters and more extensive simulations are needed in order to cover a large range of scenarios and possible outliers.
- The underlying path-loss model (i.e., the simplified path-loss model of equation (4.2)) may seem over-simplistic (without taking into account the geometry of the buildings). However, this simplified path-loss model has been well supported by the field measurement data indoors and it currently gives the best trade-off between simplicity (e.g., number of involved parameters) and accuracy (e.g., fit to real data scenarios).
- The MATLAB simulator developed until now is without some graphical user interface (GUI). Thus, it is not yet very user friendly for everyone in order to carry out the simulation. However, the purpose of the simulator so far has been in helping project-related research work and cannot be used currently for teaching activities or as an open-source model, thus we believe that a GUI will bring no added value at this point.

6. Comparison between measurement-based and simulation-based results

In this chapter, the simulation results based on the MATLAB simulator introduced in Chapter 5 are presented. The results shown here are partly from the simulated data and partly from the real measured data. The reason to show also some results based on real data measurements is for comparison purposes (in order to compare the MATLAB simulator with results achieved from field measurements).

All the real measured data used here are for indoor WLAN networks and the simulations have been carried out considering the parameters for WLAN networks. The observations which cannot be carried out with the real measured data due to lack of real measured data or due to the lack of all required information in the measured data have been carried out with the aid of the simulated data in MATLAB. The simulator was first developed for 2D model and then extended to 3D, hence most of the observation results are for 2D estimation and only few for 3D estimation.

6.1. Observations/results based on real measured data

6.1.1. AP position estimation error

This section addresses the problem of the accuracy of the estimation of the APs' location (since AP location is typically unknown). The data used in this section is based on field measurement data gathered from one building of Tampere University of Technology. In that particular building, the position of 16 APs were known. The data has been collected separately for two floors using a Windows tablet with incorporated WLAN receiver and associated software. The building map was also known. Since the map used in the software incorporated with the tablet only gave the (x, y) location of the reference point, the errors presented here are only for 2D. If a physical AP location had several MAC addresses associated to it (e.g., multiple-input and multiple-output (MIMO) WLANs or multiple SSID), then the different MAC addresses coming from the same AP were considered as different APs. *Table 6.1*shows the estimation error in x and y coordinate when the same floor's database has been used. This example is for an AP with 4 associated MAC addresses, coming from the same physical location. That is the reason for getting 4

estimated coordinates pairs for a single true position. The AP showed in this example was located at the first floor. Since the AP was heard promptly even in the second floor, the AP could also be estimated using the adjacent floor's database and the estimation error using the second floor's database has been shown in the *Table 6.2*.

Table 6.1 *Error in APs estimation using same floor's fingerprint data.*

Using the dataset from 1st floor only			
True position (x_{true}, y_{true})	**Estimated positions** (x_{est}, y_{est})	**Estimated error** (x_{error}, y_{error})	**Estimated error (distance m)**
(-67.59, 20.53)	(-66.96, 20.81)	(0.63, 0.28)	0.68
	(-67.03, 20.84)	(0.56, 0.31)	0.64
	(-66.65, 20.63)	(0.94, 0.10)	0.94
	(-66.74, 20.37)	(0.85, 0.16)	0.86

Table 6.2 *Error in APs estimation using adjacent floor's fingerprint data.*

Using the dataset from 2nd floor only			
True position (x_{true}, y_{true})	**Estimated position** (x_{est}, y_{est})	**Estimated error** (x_{error}, y_{error})	**Estimated error (distance m)**
(-67.59, 20.53)	(-67.41, 21.87)	(0.18, 1.34)	1.35
	(-67.57, 22.30)	(0.02, 1.77)	1.77
	(-67.50, 22.06)	(0.09, 1.53)	1.53
	(-67.41, 21.83)	(0.18, 1.30)	1.31

From the above tables, the estimation of APs seems to be quite accurate with less than 1 m error using the same floor's dataset and less than 2 m error when using the adjacent next floor's dataset. The above tables also illustrate the concept of virtual APs, where an AP is in fact heard over multiple floors. If 2D model is used (floor-by-floor model), then the same APs will be estimated several times, as a 'virtual AP', on each floor where it is heard.

An average over all the 16 known AP locations in one of the buildings of Tampere University of Technology is illustrated in *Table 6.3*.

Table 6.3 *Path-loss parameters estimation for all known APs in TUT building.*

	Est. average path-loss coefficient $\tilde{n}$	**Est AP average transmitted power** $\tilde{P}_{T,ap}$ **[dBm]**	**RMSE for AP location [m]**
1st floor AP, 1st floor data	3.44	-25.98	1.50
1st floor AP, 2nd floor data	2.98	-42.20	2.23
2nd floor AP, 1st floor data	2.19	-49.45	24.40
2nd floor AP, 2nd floor data	3.98	-12.69	6.04

If the measurements were in 3D (x,y,z) then there might have been some greater error while using the adjacent floor's dataset but this is quite not visible with 2D model.

The above example is one of the best-cases results, but there were few cases when there was even 40-50m of error in estimation. Those errors were mostly found in the environment where we have open space no demarcation between the two floors.

It was not possible to find the errors in the estimation of path-loss exponent n and the apparent transmit power P_T since we do not know them in real measurements. This has been observed with some simulated data presented later in this chapter.

6.1.2. MS position estimation error

Here, we observed the accuracy of MS estimation from the collected data in TUT. We had around 500 fingerprint data and a user track containing 15-18 user points in 150m by 100m floor containing many office rooms, lecture rooms and long open corridors. From those measured data, MS estimation was done using fingerprinting method and path-loss methods.

In both the methods, Euclidian distance approach has been used in order to find the difference in RSS between the MS and the fingerprints and the estimation is done averaging over 5 NN. In the path-loss model we reconstructed the grid of step size 2m by 2m, 10m by 10m and 40m by 40m. At some user points the amounts of hearable APs were quite large, 50 or even more. So, it might not be relevant to keep all the hearable APs in the positioning phase [10]. Hence, we have also limited the number of hearable APs (limit numbers are

shown in the table below). The APs are sorted in descending order based on the RSS at the MS and then only limited number of them is considered. The RMSE in each of the cases have been presented below in *Table 6.4*.

Table 6.4 *RMSE in user estimation using fingerprinting and path-loss method, measurement data, 2D model.*

		Path-loss			**Fingerprinting**
Scenario	**No. of hearable APs limited to**	**RMSE [m] for grid size**			**RMSE [m]**
		2m by 2m	**10m by 10m**	**40m by 40m**	
TUT building	3	8.01	6.95	8.16	7.09
	5	6.91	6.24	9.36	8.56
	7	5.62	5.70	10.75	6.26
	10	5.69	6.56	11.86	6.01
	15	6.86	7.60	12.18	6.49
	all	12.96	13.41	11.80	7.52

In most of the cases, the RMSE of MS position estimation error is within 10m. Both the fingerprinting and path-loss model give almost the same range of error. In the path-loss model, the small enough grid step does not better the approximation error as the information contained is almost the same even if we have 2m grid or 10m grid size (because, the RSS levels are reconstructed based on the simplified path-loss model). Some information is lost when the grid size is large enough, such as 40 m x 40 m grid. From the above table, it is seen that the errors using 2m or 10m grid size are more or less same, but there is around 5m increase in error when the grid size is increased to 40m. A typical grid size seems therefore to be in the range of 2 x 2 till 10 x 10 m (smaller grid would also increase the computational complexity).

*Figure 6.1*shows the pictorial representation of the estimation of user track. The plot shows the comparison between the true user track (in black), user-track estimated via fingerprinting method (in green) and path-loss model using 10m by 10m grid size (in red).

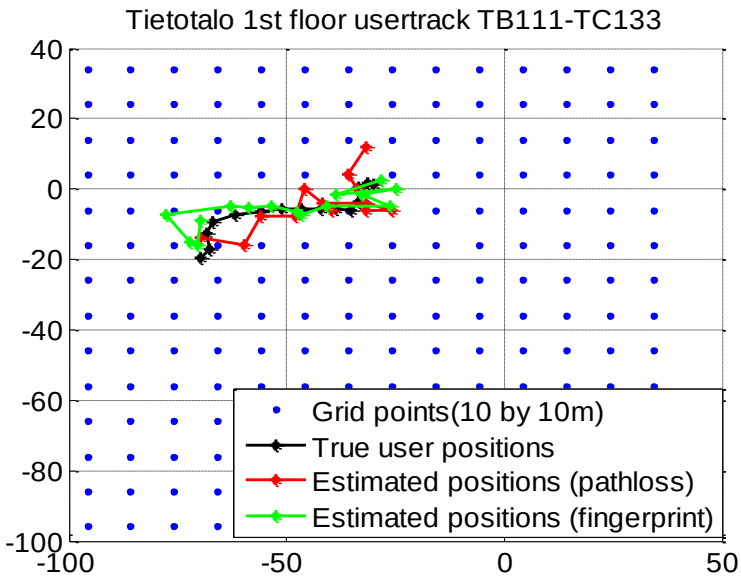

Figure 6.1 *True and estimated user-track.*

6.1.3. Average floor attenuation

A parameter of interest in building our simulator is the average floor attenuation from one floor to another in multi-floor buildings. The analysis has been done on four buildings in Tampere, and an example is shown below in *Figure 6.2*, for a 4-floor shopping center building.

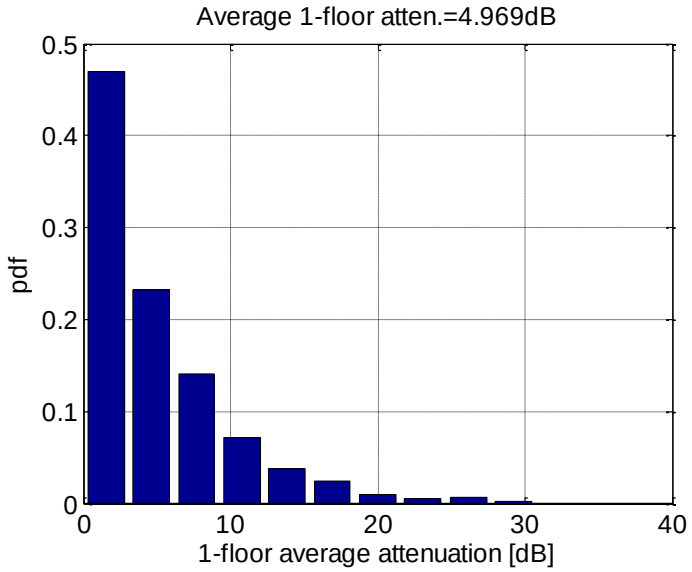

Figure 6.2 *Estimated floor attenuation based on a multi-floor shopping center in Tampere.*

The floor attenuation was estimated by comparing average RSS per AP at successive floors. As seen in *Figure 6.2*, the floor attenuation is typically below 10 dB. The considered shopping center was having a special topology with open spaces between floors, which explain why the average floor attenuation (over all heard APs on more than one floor) is as small as 1 dB in almost 45% of the cases.

Another situation, this time in the University building is illustrated in *Figure 6.3* below.

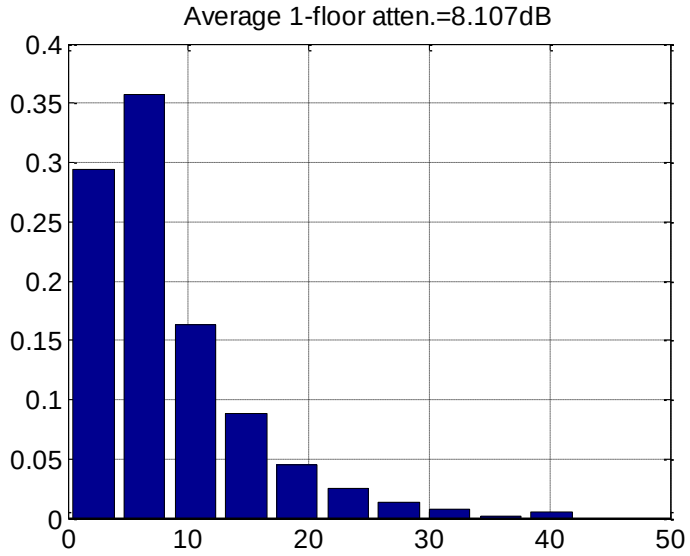

Figure 6.3 *Estimated floor attenuation based on a multi-stored building in (TUT).*

The average floor attenuation for concrete walls, as those from university is about 8 dB, while the average floor attenuation in malls /shopping centers with open spaces between floors is about 6 dB. Such values, as inferred from the above plots, have also been used in our MATLAB simulator, where the floor attenuation was assumed to be uniformly distributed between 5 and 35 dB (with equal average floor attenuation per building or scenario).

6.1.4. Various metrics to choose the nearest neighbor

In the above observation, Euclidian power difference has been used to find the difference in RSS between the MS and the fingerprints. Euclidian metric and some other metrics studied so far in the thesis are as follows:

Euclidian distance

The combined power difference at the i^{th} measurement point is calculated as:

$$\Delta P_i = \sum_{ap=1}^{N_{ap}} \left(P_{ap}^{(i)} - P_{ap}^{(MS)} \right)^2 \tag{6.1}$$

or

$$\Delta P_i = \prod_{ap=1}^{N_{ap}} \left(P_{ap}^{(i)} - P_{ap}^{(MS)} \right)^2 \tag{6.2}$$

for i=1,…..N measurement point index

where, ΔP_i is the combined power difference at the i^{th} measurement point for ap =1,.. N_{ap} hearable APs index, $P_{ap}^{(i)}$ is the RSS at i^{th} measurement point and $P_{ap}^{(MS)}$ is the RSS at MS/user point due to ap^{th} AP.

The combined power difference at each of fingerprint point (i.e. i=1,…..N No. of fingerprints or grid points in the reconstructed path-loss grid) is calculated and then minimized. The measurement point/points with the least power difference are considered as the closest match.

Manhattan Distance

The combined power difference at the i^{th} measurement point is calculated as:

$$\Delta P_i = \sum_{ap=1}^{N_{ap}} abs\left(P_{ap}^{(i)} - P_{ap}^{(MS)} \right) \tag{6.3}$$

or

$$\Delta P_i = \prod_{ap=1}^{N_{ap}} abs\left(P_{ap}^{(i)} - P_{ap}^{(MS)} \right) \tag{6.4}$$

for i=1,…..N measurement point index

where, ΔP_i is the combined power difference at the i^{th} measurement point for ap =1,.. N_{ap} hearable APs index, $P_{ap}^{(i)}$ is the RSS at i^{th} measurement point due to ap^{th} AP and $P_{ap}^{(MS)}$ is the RSS at MS/user point due to ap^{th} AP.

The combined power difference at each of fingerprint point is calculated and is minimized and the measurement point/points with the least power difference are considered as the closest match as in Euclidian method.

Entropy based

In the entropy based method, the combined power difference is calculated as:

$$z_{ap}^{(i)} = abs\left(P_{ap}^{(i)} - P_{ap}^{(MS)}\right) \tag{6.5}$$

$$\Delta P_i = \sum_{ap=1}^{N_{ap}} z_{ap}^{(i)} * log_2\left(z_{ap}^{(i)}\right) \tag{6.6}$$

or

$$\Delta P_i = \prod_{ap=1}^{N_{ap}} z_{ap}^{(i)} * log_2\left(z_{ap}^{(i)}\right) \tag{6.7}$$

for i=1,…..N measurement point index

where, $P_{ap}^{(i)}$ is the RSS at i^{th} measurement point and $P_{ap}^{(MS)}$ is the RSS at MS/user point due to ap^{th} AP, $z_{ap}^{(i)}$ is the absolute RSS level difference at i^{th} measurement point due to ap^{th} AP and ΔP_i is the entropy based combined power difference at i^{th} measurement point for ap =1,.. N_{ap} hearable APs index.

The combined power difference at each of fingerprint point is calculated and then the power difference is maximized this time. The measurement point/points with the largest power difference are considered as the closest match.

Weighted Average

Here, we directly find the MS location as the weighted average of the location of APs it can hear and the measured RSS from those APs. The location of MS is calculated as weighted average as:

$$MS(x,y) = \left(\frac{\sum_{ap=1}^{N_{ap}} x_{ap}\ P_{ap}^{(MS)}}{\sum_{ap=1}^{N_{ap}} P_{ap}^{(MS)}} , \frac{\sum_{ap=1}^{N_{ap}} y_{ap}\ P_{ap}^{(MS)}}{\sum_{ap=1}^{N_{ap}} P_{ap}^{(MS)}} \right) \quad (6.8)$$

Where, (x_{ap}, y_{ap}) is the location of ap^{th} AP and $P_{ap}^{(MS)}$ is the RSS at MS from ap^{th} AP and the average is over N_{ap} hearable APs at the MS. The estimation has been observed keeping the RSS at the MS i.e. $P_{ap}^{(MS)}$ both in linear scale and logarithmic scale.

Likelihood with shadowing

In this metric, the standard deviation and mean shadowing from each AP are also considered to find the difference in RSS levels of the MS and the fingerprints. The combined power difference is calculated as:

$$\Delta P_i = \sum_{ap=1}^{N_{ap}} \frac{1}{2\pi\sigma_{ap}} e^{\left(\frac{\left(P_{ap}^{(MS)} - \mu_{ap}\right)^2}{2\sigma_{ap}^2} \right)} \quad (6.9)$$

or

$$\Delta P_i = \prod_{ap=1}^{N_{ap}} \frac{1}{2\pi\sigma_{ap}} e^{\left(\frac{\left(P_{ap}^{(MS)} - \mu_{ap}\right)^2}{2\sigma_{ap}^2} \right)} \quad (6.10)$$

for i=1,…..N measurement point index

where, ΔP_i is the combined power difference at the i^{th} measurement point, $P_{ap}^{(MS)}$ is the RSS at MS/user point due to ap^{th} AP and σ_{ap} and μ_{ap} are the standard deviation and mean of shadowing of ap^{th} AP, respectively. The combined power difference is either sum/product of all hearable APs. Hence the fingerprint/fingerprints with the minimum power difference is/are considered as the closest match.

Using the above mentioned metrics the estimations were carried out and the estimation RMSE in different scenarios have been mentioned below in *Table 6.5* and *Table 6.6*:

Table 6.5 *RMSE with fingerprinting method using various metrics to find NN.*

Scenarios	RMSE [m]					
	Euclidian		Manhattan		Entropy Based	
	Sum	product	sum	product	sum	product
Tietotalo1 (TUT)	6.42	6.97	6.98	6.97	20.02	20.13
Tietotalo2 (TUT)	13.22	14.09	13.15	14.09	37.68	37.34
Sähkotalo1 (TUT)	11.27	11.54	11.46	11.54	29.52	29.03
Sähkotalo2 (TUT)	9.97	10.59	9.95	10.59	22.07	22.17
Shopping center #1	24.37	25.16	24.50	25.16	30.56	29.40
Shopping center #2	11.55	13.54	11.53	13.54	39.96	38.14
Shopping center #3	11.96	20.68	12.11	20.68	27.72	27.90

Table 6.6 *RMSE with path-loss method using various metrics to find NN.*

Scenarios	RMSE [m]									
	Euclidian		Manhattan		Entropy Based		Weighted Avg.		Likelihood with shadowing	
	sum	product	sum	product	sum	product	linear	log	sum	product
Tietotalo 1 (TUT)	9.41	9.68	10.14	9.68	134.09	144.68	10.65	8.85	144.69	144.68
Tietotalo 2 (TUT)	12.17	15.55	12.34	15.55	138.38	135.22	14.22	17.56	123.90	138.40
Sähkotalo 1 (TUT)	15.16	15.60	14.15	15.60	176.12	176.10	14.28	13.95	172.27	174.83
Sähkotalo 2 (TUT)	9.57	13.09	10.03	13.09	172.44	174.03	9.92	12.04	165.72	171.47
Shopping center #1	28.53	30.29	27.90	30.29	115.50	113.44	29.86	24.75	110.40	103.91
Shopping center #2	15.30	21.98	16.58	21.98	168.78	171.56	20.72	21.39	167.76	167.27
Shopping center #3	14.11	19.94	13.17	19.94	71.00	72.23	16.74	22.09	56.80	67.15

In *Table 6.5* and *Table 6.6* the statistics have been done over 15 - 100 user points, according to the available measured tracks. The RMSE has been computed with limiting the number of hearable APs to 10 and the averaging is done over 5 NNs, except in weighted average metric which directly gives the MS position. *Table 6.5* gives the RMSE with fingerprinting method and RMSE with path-loss model is given in *Table 6.6*. More metrics have been investigated in path-loss method than in fingerprinting method. Those methods

are not suitable for fingerprinting method as they require some path-loss parameters like APs' location, standard deviation and mean shadowing of APs. From the tables, Euclidian distance, Manhattan distance and weighted average metrics are giving better results than other metrics. Euclidian metric using the summation for finding the combined power is the best among all other metrics for most of the cases and therefore it will be used in what follows.

6.2. Observations based on simulated data

6.2.1. Simulated model parameters

The main parameters used in our MATLAB model are summarized in the *Table 6.7*below

Table 6.7 *Main parameters of the MATLAB simulator.*

Parameter	**Value/Ranges**
Floors per building in one scenario	4
Number of random scenarios (observations)	200
Shadowing standard deviation	Fixed values between 0 and 10 dB (specified in the results below)
Transmit powers of APs	Uniformly distributed between -20 and 0 dBm (constant per AP)
Minimum hearable RSS	-100 dBm
Building sizes	Uniform distributions for x and y directions between 50 and 200m; equal floor height of 4 m
Total number of fingerprints per scenario	Uniformly distributed between 400 and 1000
Multiple SSID percentage	50%
Multiple MAC addresses	Between 3 and 5
Total number of APs per building	Uniformly distributed between 50 and 200
Path-loss coefficient	Uniformly distributed between 1.2 and 10
Minimum hearable distance	Uniformly distributed between 6 and 20 m
Maximum hearable distance	Uniformly distributed between 35 and 60 m
AP distribution per floor	Unequally distributed, scenario dependent
Floor attenuation	Uniformly distributed between 5 and 35 dB

6.2.2. AP location estimation error

A snapshot of the AP distribution in our MATLAB model with 4-floor buildings is illustrated in *Figure 6.4*.

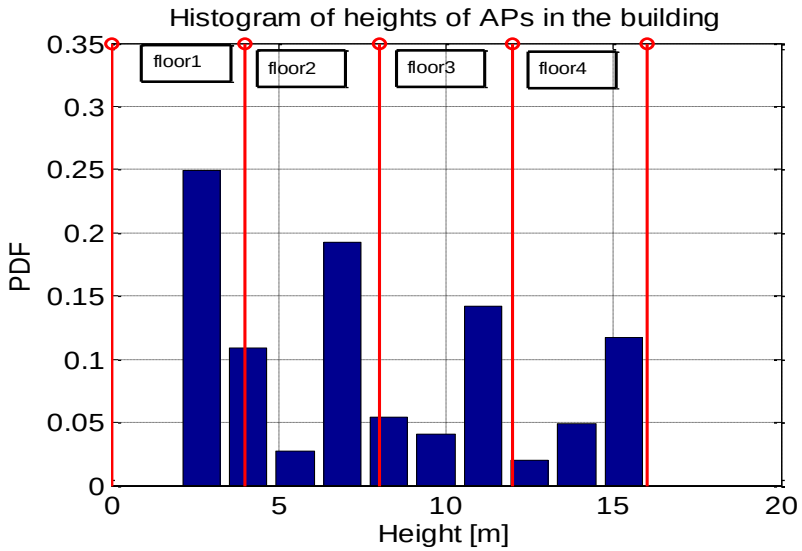

Figure 6.4 *Example scenario of APs distribution in MATLAB model in 3D case.*

The floor heights in this example were 0, 4, 8 and 12 m. The estimated AP location error with the 3D model has been 3.95 m (x-y-z error, averaging over 200 different scenarios), which is very good and also close to what has been observed from measured data. The simulated results are typically slightly worse than the results based on the measured data because we capture much more variations of the environment with our simulated model.

6.2.3. Path-loss parameters estimation error

In the real measurements we do not know the path-loss exponent, apparent transmit power and in most of the cases APs location, too. But in the simulated data, we know these parameters beforehand. In this section we investigate how accurate is the simulator in terms estimating the path-loss parameters. The following table gives the minimum error, maximum error and RMSE in the estimating the APs location, path-loss exponent and the apparent transmit power. The calculations here has been done over 400 APs randomly located in 100m by 100m by 16m building and each AP was associated P_T randomly varying between -40 to 0 dBm and with n=5. The RMSE has been calculated as:

$$RMSE = \sqrt{\frac{1}{N_{ap}} \sum_{i=1}^{N_{ap}} (error_i)^2} \tag{6.11}$$

where,

For the path-loss exponent case,

$error_i = (n_{true} - n_{est})_i$ = true path-loss exponent minus estimated path-loss exponent for i^{th} AP.

For apparent transmit power case,

$error_i = (P_{T\,true} - P_{T\,est})_i$, = true transmit power minus estimated transmit power for i^{th} AP.

For the AP location case,

$error_i = (\sqrt{(x_{true} - x_{est})^2 + (y_{true} - y_{est})^2 + (z_{true} - z_{est})^2})_i$ = true AP location minus estimated AP location for i^{th} AP.

The averaging for the error is done over i = 1,.. N_{ap} AP index.

The APs have been estimated with an error less than 1 m in minimum, around 17 m in maximum and on average with 6 m error. The mean errors in estimating the path-loss exponent and the apparent transmit power are 1.05 and 23.59 dBm, respectively. We will see later than under-estimating or over-estimating the apparent transmit power of the AP does not have a great impact on the accuracy of MS position estimation. We believe that this is due to the fact that there is also a shadowing effect in our model, and therefore the apparent transmit power may be influenced by the shadowing level.

Table 6.8 *Estimation error for path-loss parameters.*

AP location [m]			**Path-loss Exponent *n***			**Apparent Transmit power P_T [dBm]**		
Min Error	Max Error	RMSE	Min Error	Max Error	RMSE	Min Error	Max Error	RMSE
0.82	16.94	5.74	0.03	2.95	1.05	0.16	57.12	23.59

6.2.4. Multiple Nearest Neighbors (NNs)

In here we investigate how the number of nearest neighbors used in the averaging are influencing the MS location estimation Here, the basic analysis has been presented taking the NN from 1 to 7 and using the fingerprinting estimation method.

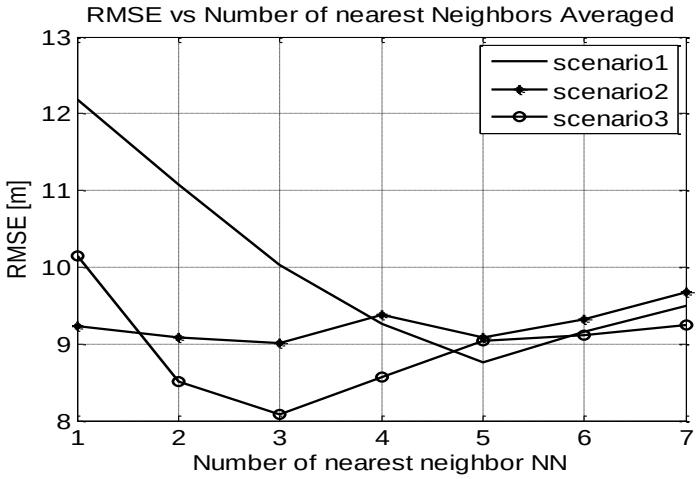

Figure 6.5 *Estimation RMSE with varying nearest neighbors*.

In *Figure 6.5*, the MS estimation error for 3 different scenarios has been presented. The RMSE is the average over 25-50 user points. It is seen that multiple NNs are always better than simply picking only one closest neighbor. It is likely that the error vector corresponding to each neighbor is oriented in different directions and averaging the coordinates of those neighbors may yield closer user estimates. Often there might be multiple neighbors that are roughly at same distance from the true user position, so, it is better to consider all of them instead of rejecting them [5].

6.2.5. Estimation error with varying shadowing standard deviation

In this section we investigate how the positioning error varies with the varying shadowing standard deviation. The simulation was carried over 500 scenarios with each scenario containing 120 user (or track) points. In each scenario the location of user points, fingerprint data, location of APs were varied randomly, typically according to uniform distributions, with bounds selected in accordance with the field measurement data. The fingerprinting data was generated varying the added uncorrelated shadowing standard deviation as 4dB, 8dB, 12dB, 16dB and 20dB and the estimation error in each case were recorded. The result presented here in *Figure 6.6* is the RMSE based on distance in 3D for both fingerprinting and path-loss methods. The plot clearly shows that the location estimation accuracy decreases in the environment with large shadowing deviation, as

expected. However, the slope of decreased accuracy level is not very high and the RMSE still remains below 15 m even at 20 dB shadowing standard deviation.

Figure 6.6 *Estimation RMSE with varying shadowing deviation.*

6.2.6. Bias in the estimation phase compared to training phase

In this section we test what happens when we have two different devices to perform the training and the estimation phases. For example, the training data is collected with a Windows tablet and the estimation is done with a mobile phone. This translates to the fact that here will be some bias in the RSS levels from the estimation phase compared to those coming from the training phase. While generating the data, the user points were given some $+B$ and $-B$ bias (in dB) in RSS. The plots in *Figure 6.7* below show how much the RMSE in biased estimation deviates from the RMSE in without biased estimation when we have +/-5dB and +/-10dB biases in the RSS in the user device.

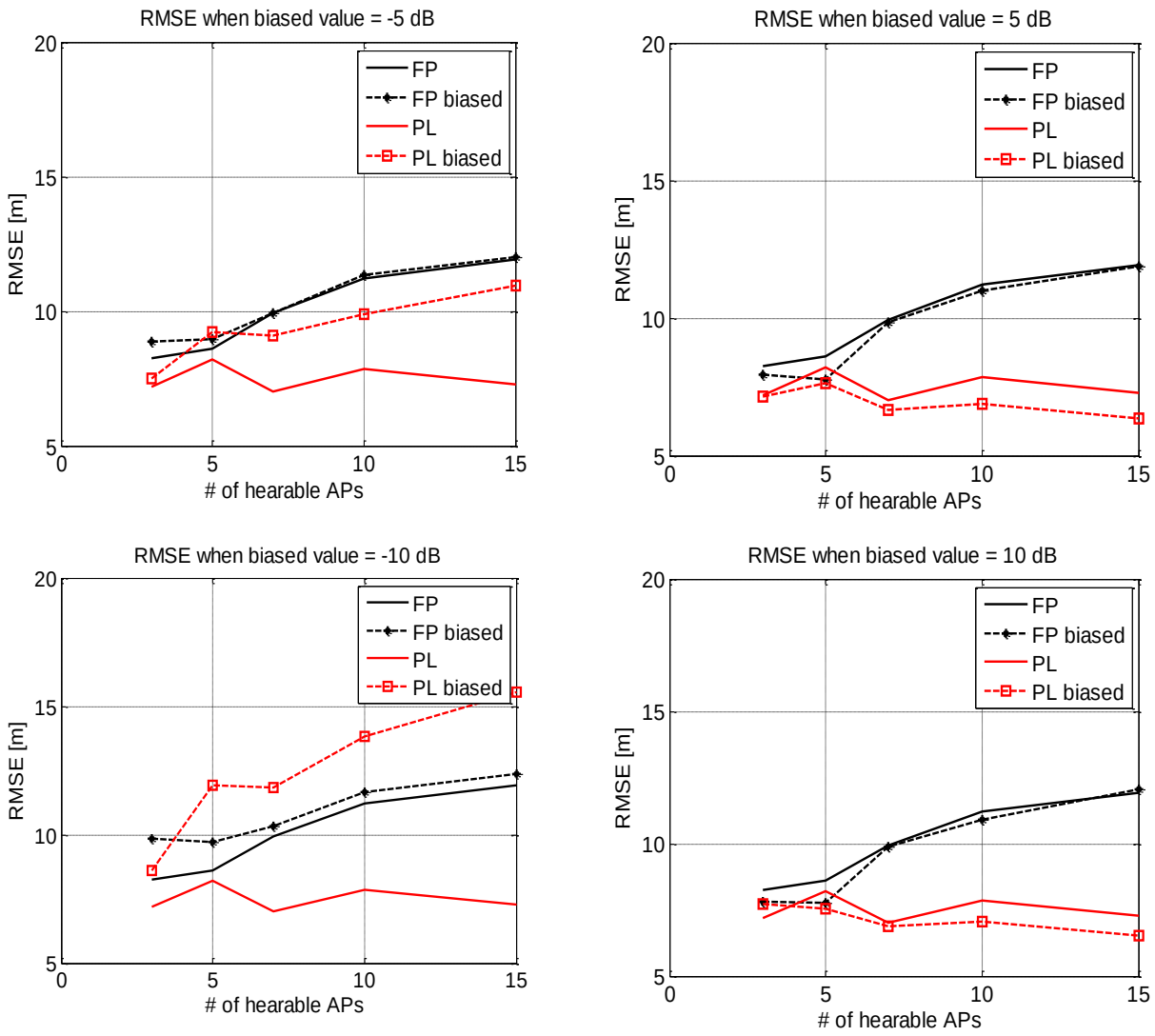

Figure 6.7 *Estimation error with biased RSS in the user points*

The bias value is visible more in path-loss model than in fingerprinting method. A negative bias seems to have larger effect in estimation than a positive bias, which is intuitively true because a positive bias is seen as the MS being closer to AP that it apparently is, and a shorter MS-AP distance usually means better accuracy in the estimation.

7. Conclusions and future works

7.1. Conclusions

The main contributions of this thesis have been to analyze several RSS-based positioning methods for cellular (2G and 3G) and WLAN signals and to develop a unique and novel MATLAB simulator for indoor positioning in multi-floor buildings based on RSS methods. The methods used in this simulator were both fingerprinting methods (based on fingerprinting and various nearest neighbor metrics) and path-loss methods. Several of the results shown in this thesis especially for the radio channel measurements were based on the true measurements. Those measurements campaigns were carried out by the Author together with the other project team members for one of the ongoing projects in the Department of Communication Engineering (DCE) at TUT. The parameters derived based on real channel measurements were later used as adjustable parameters in the MATLAB simulator developed in this thesis work.

Based on the measured data, the Author studied the characteristics of RSS in different environments, by paying a special attention to path-losses and shadowing models. The RSS fluctuations were studied in terms of radio maps, shadowing distributions, shadowing de-correlation distances, apparent transmit powers of APs/BSs, path-loss exponents, and floor attenuations in multi-floor buildings. The study was carried out for both cellular and WLAN cases both indoor and outdoor scenarios. Based on this detailed study, conclusions were made regarding the range of shadowing, shadowing de-correlation distance, path-loss exponent values in the indoor and outdoor cases and these conclusions were used as the basis to build the indoor 2D and 3D positioning models. A high similarity regarding path-loss and shadowing models has been observed between the cellular and WLAN signals, despite their operations at different carrier frequencies. Typically, WLAN signals exhibited higher standard deviations of shadowing than cellular signals.

Regarding the simulation results, the MATLAB simulator gave more or less similar accuracy as in the real-time positioning. But it can never be overlooked that the positioning accuracy largely depends on the environment as the RSS database is environment dependent. The simulator was also tested for estimating the APs location and the path-loss parameters and it also gave quite accurate results. The simulator was also tested with varying environment conditions, such as varying shadowing standard deviation and varying

path-loss coefficients. Hence, the simulator can become quite an easy tool to do extensive study in RSS-based positioning which is often hard to do with real measurements.

7.2. Future work

There is still room for improvement in this thesis. The work will be continued by making the MATLAB simulator even more generic and by overcoming the shortcomings enumerated in section 5.3. The way it currently models the environment by taking into account only some shadowing, the floor attenuation and hear ability threshold as described in section 5.1 is still based on state-of-art algorithms at the moment. For example, the simulator cannot model the true building topology such as walls, rooms, furniture etc. Therefore, in the future it can be enhanced in some way that it may read a map and create the RSS database according to it. However, a more complex model with more parameters may not be necessarily better suited for large-scale indoor positioning and this is a research topic that remains open in the current literature. The simulator provides the user estimation using the fingerprinting and the state-of-art one-parameter path-loss model. Other path-loss models and other estimating approaches can be further developed and implemented which may represent a topic of further research.

References

[1] P. Prasithsangaree, P. Krishnamurthty, P.K. Chrysanthis, "On Indoor Position Location with Wireless LANs", *13th IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications,* PIMRC 2002.

[2] Y Wang, X. Jia, H.K. Lee, "An indoor wireless positioning system based on wireless local area network infrastructure", presented in *6th International Symposium on Satellite Navigation Technology Including Mobile Positioning & Location Services,* Melbourne, Australia, 22–25 July 2003.

[3] GSA GNSS Market report – Issue 1, Oct 2010.

[4] Shweta Shrestha, Elina Laitinen, Jukka Talvitie and Elena Simona Lohan, "RSSI channel effects in cellular and WLAN positioning", accepted at *the 9th Workshop on Positioning, Navigation and Communication (WPNC),* Mar 2012.

[5] Paramvir Bahl, Venkata N. Padmanabhan, "User location and tracking in an in – building radio network", *Technical Report MSR-TR-99-12, Microsoft research,* Microsoft Corporation, February 1999.

[6] Carl R. Stevenson, WK3C Wireless Gerald Chouinard, Communications Research Centre, Canada Zhongding Lei, Institute for Infocomm Research, Singapore Wendong Hu, STMicroelectronics, Inc.Stephen J. Shellhammer, Qualcomm Inc.Winston Caldwell, Fox Technology Group, "IEEE 802.22: The First Cognitive Radio Wireless Regional Area Network Standard", *IEEE Communications Magazine* , January 2009.

[7] Dragorad A. Milovanovic , K. R. Rao , and Zoran S. Bojkovic, "Wireless Multimedia Communications", CRC Press , Pages 99–166, 2008.

[8] Stanislav Safaric, Kresimir Malaric, "ZigBee wireless standard", *48th International Symposium ELMAR-2006,* Zadar, Croatia, 07-09 June 2006.

[9] Justin Thiel, "Metropolitan and Regional Wireless Networking: 802.16, 802.20 and 802.22", 2006. www.cs.wustl.edu/~jain/cse574-06/ftp/wimax.pdf.

[10] Amit Kumar, Dr. Yunfei Liu, Dr. Jyotsna Sengupta, Divya, "Evolution of Mobile Wireless Communication Networks:1G to 4G", *68 International Journal of Electronics & Communication technology,* IJECT Vol. 1, Issue 1, December 2010.

[11] Elina Laitinen, Elena Simona Lohan, Jukka Talvitie, Shweta Shrestha, "Access point significance Measures in WLAN- based location", *accepted at the 9th Workshop on Positioning, Navigation and Communication (WPNC),* Mar 2012.

[12] Theodore S. Rappaport, "Wireless Communications Principles and Practice", second edition, 2002.

[13] S. Saunders, "Antennas and Propagation for WirelessCommunication Systems", Wiley, 2000, 409 p.

[14] Ingo Forkel, Marc Schinnenburg, Markus Ang, "Generation of two dimensional correlated shadowing for mobile radio network simulation", *6th International Symposium on Satellite Navigation Technology Including Mobile Positioning & Location Services*, Melbourne, Australia, 22–25 July 2003.

[15] E. Cassano, F. Florio, F. De Rango, S. Marano, "A Performance Comparison between ROC-RSS and Trilateration Localization Techniques for WPAN Sensor Networks in a Real Outdoor testbed", *Wireless Telecommunications Symposium, 2009. WTS 2009*, pp-1-8, April 2009.

[16] Wonsun Bong and Yong C. Kim ,"Indoor Localization for Wi-Fi Devices by Cross-Monitoring AP and Weighted Triangulation", *in Consumer Communications and Networking Conference (CCNC), 2011 IEEE*, pp 933 – 936, Jan 2011.

[17] Bensky Alan, "Wireless Positioning Technologies and Applications", Artech house, 2007

[18] Yu Zhou, "An efficient Least- Squares trilateration algorithms for mobile robot localization*", in IEEE/ RSJ International conference on Intelligent Robots and Systems,* St.Louis, USA, 11-15 October 2009.

[19] Binghao Li, James Salter, Andrew G. Dempster and Chris Rizos, "Indoor Positioning TechniquesBased on Wireless LAN", *in First IEEE International Conference on Wireless Broadband and Ultra Wideband Communications*, www.gmat.unsw.edu.au/snap/publications/lib_etal2006a.pdf

[20] Kamol Kaemarungsi, "Distribution of WLAN Received Signal Strength Indication for Indoor Location Determination", *in Wireless Pervasive Computing, 2006 1st International Symposium,* pp 6, 16 – 18 Jan 2006.

[21] M.A.Youssef, "HORUS: A WLAN- based indoor location determination system," *Ph.D. dissertation*, University of Maryland, College Park, MD, 2004.

[22] Kamol Kaemarungsi , "Design of Indoor Positioning Systems Based On Location Fingerprinting Technique", *Graduate Faculty of the School of Information Science University of Pittsburgh, Doctor of philosophy thesis,* 2005.

[23] Antonia Kalaitzi, "Measurement-based Multipath Characterization for Indoor GPS Channels", *Msc Thesis at Tampere University of technology*, May 2009, http://www.cs.tut.fi/tlt/pos/Antonia_MSc_Apr09.pdf

[24] Mazuelas, S., Bahillo, A., Lorenzo, R.M., Fernandez, P., Lago, Francisco A., Garcia, E., Blas, J., Abril, E.J. "Robust Indoor Positioning Provided by Real-Time RSS Values in Unmodified WLAN Networks", *IEEE Journal of Selected Topics in Signal Processing, vol. 3, issue 5*, pp. 821-831, Oct 2009.

[25] Shih-Hau Fang, Tsungnan Lin, "Principal Component Localization in Indoor WLAN Environments", in *Mobile Computing, IEEE Transactions* , volume 11, issue 1, pp. 100 – 110, 2012.

[26] A. Krishnakumar and P. Krishnan, "On the Accuracy of Signal Strength-Based Estimation Techniques," *Proc. IEEE INFOCOM*, vol. 1, pp. 642-650, 2005.

[27] Al-Ahmadi, A.S., Rahman, T.A., Kamarudin, M.R., Jamaluddin, M.H., Omer, A.I, "Single-Phase Wireless LAN Based Multi-floor Indoor Location Determination System", in *IEEE 17th International Conference on Parallel and Distributed Systems (ICPADS)*, pp. 1057-1062, 2011.

[28] El-Kafrawy, K., Youssef, M., El-Keyi, A., Naguib, A., "Propagation Modeling for Accurate Indoor WLAN RSS-Based Localization", *IEEE 72nd Vehicular Technology Conference Fall (VTC 2010-Fall)*, 2010.

[29] M. Cypriani, F. Lassabe, P. Canalda, and F. Spies, "Open wireless positioning system: a Wi-Fi-based indoor positioning system," *in Vehicular Technology Conference (VTC)*, Fall 2009.

[30] Yubin Xu, Yong Wang, Lin MaA, "Novel WLAN Indoor Positioning Algorithm Based on Positioning Characteristics Extraction", *Fourth International Conference on Genetic and Evolutionary Computing (ICGEC)*, 2010.